VERIFICATION OF TRANSLATION

I, Minoru KUDOH
of a citizen of Japan residing at: 406, 17-15,
Minamiooi 1-chome, Shinagawa-ku, Tokyo 140, Japan
certify that I am familiar with the English and Japanese languages, and to the best of my knowledge and belief the following is a true translation of the Japanese Patent Application No. 2000-003041.

This 10 day of November, 2005

Minoru KUDOH

[Document Name] PATENT APPLICATION
[Identification No.] 49210398
[Filing Date] January 11, 2000
[To] Commissioner; Japanese Patent Office
[International Patent Classification] H04L 12/56
[Inventor]
[Domicile or Residence] c/o NEC Corporation, 7-1, Shiba 5-chome, Minato-ku, Tokyo, Japan
[Name] MASUDA Michio
[Inventor]
[Domicile or Residence] c/o NEC Corporation, 7-1, Shiba 5-chome, Minato-ku, Tokyo, Japan
[Name] ARIKAWA Toshiaki
[Inventor]
[Domicile or Residence] c/o NEC Corporation, 7-1, Shiba 5-chome, Minato-ku, Tokyo, Japan
[Name] YAMADA Kenshin
[Applicant]
[ID number] 000004237
[Name] NEC Corporation
[Representative] KISHI Sadayuki
[Attorney]
[ID number] 100065385
[Name] YAMASHITA Johei
[Telephone] 03-3431-1831
[Indication of Charge]
[Deposit Payment Register Number] 010700
[Amount of Fee] 21000
[Items of the Filing Articles]
[Article Name] Specification one copy
[Article Name] Drawings one copy
[Article Name] Abstract one copy
[General Power of Attorney] 9803702

[Document Name] Specification

[Title of the Invention] MULTI-LAYER CLASS IDENTIFYING COMMUNICATION APPARATUS AND COMMUNICATION APPARATUS

[Scope of the Invention to be Claimed]

[Claim 1]

A multiplayer class identifying communication apparatus comprising an input interface, a switch circuit and an output interface and capable of selecting a plurality of levels of communication quality of communication lines, comprising:

means for determining a class identifier indicative of one of classes to which an IP packet belongs, from a header data of said IP packet received said input interface connected to said communication network (a combination of a layer 3 data represented by an IP header and a layer 4 data represented by a TCP/UDP header corresponding to a higher level layer), and allocating an IP-QoS (Internet Protocol Quality of Service) code in handling each of the IP packet processes in said input interface to said IP packet, and

wherein output scheduling is carried out for each of layers of a layer 4 or above from said output interface.

[Claim 2]

The multi-layer class identifying communication apparatus according to claim 1, comprising:

means for defining a priority traffic based on an arbitrary combination of said IP header and values of a plurality of fields of a TCP header, regarding allocation of the IP code (mapping of traffic class).

[Claim 3]

The multi-layer class identifying communication apparatus according to claim 1, comprising:

means for using WRR (Weighted Round Robin Scheduling) and a fixed priority scheduling, for permitting to select each class with a fixed priority based on said class identifier, and for permitting a minimum band designation.

[Claim 4]

The multi-layer class identifying communication apparatus according to claim 1, wherein said apparatus comprising:

said input interface;

an output interface;

a switch carrying out switching to destination address between said input interface and said output interface; and

a scheduler indicating switch timing of said switch between said input interface and said output interface,

wherein said input interface, said output interface and

said switch operates in conjunction to perform the priority control that carries out predetermined priority control by an instruction from said scheduler, based on said IP-QoS code.

[Claim 5]

The multi-layer class identifying communication apparatus according to claim 1, further comprising queue managing section which manages a queue such that a plurality of IP packets can be shared in units of said IP-QoS codes to obtain statistical multiplexing effect.

[Claim 6]

The multi-layer class identifying communication apparatus according to claim 4, wherein each of said input interface and said output interface monitors traffic in units of said IP-QoS codes to restrict excessive traffic.

[Claim 7]

The multi-layer class identifying communication apparatus according to claim 4, wherein said class identifier includes three kinds of traffic (service class) of an EF (Expedited Forwarding (Premium service)) class, an AF (Assured Forwarding Service) class, and a BE (Best Effort Service) class.

[Claim 8]

A communication apparatus comprising an input interface and an output interface in which a plurality of quality levels of communication of communication network can be selected,

wherein said input interface comprises:

an IP packet receiving section which extracts said header data and TCP header data of received IP packet;

an IP-QoS class determining section which refers to class IP-QoS code (class identifier) memory to determine said class identifier, by using said header data of said IP packet as a search key;

a reception-side switch interface control section which carries out a priority control to the IP packet that a destination has been specified, based on said IP-QoS class code and IP packet data of said IP packet; and

a reception side switch interface which carries out said priority control and issues a transmission request to said output interface card in units of said class identifiers,

said IP-QoS class determining section monitors a coming traffic which exceeds a transmission permissive capacity which is set for every IP-QoS class, carries out a discarding operation of IP packets of said coming traffic or a policing operation to lower transmission priorities of said IP packets of said coming traffic, when said coming traffic exceeds said transmission permissive.

[Claim 9]

A communication apparatus comprising an input interface and an output interface in which a plurality of quality levels of communication of communication network can be selected, comprising:

a transmission-side switch interface that said output interface which receives priority-controlled IP packets transmitted from said input interface to store in transmission-side payload memory, and generates IP packet data to write in said FIFO memory;

IP-QoS class scheduler which has a control function that carries out a scheduling function and a queuing operation based on IP-QoS class code to primarily issue a transmission request such that said IP packet is transmitted with a priority;

a transmission-side switch interface control section which transmits IP packet transferred from said input interface according to an order scheduled based on priority of said IP-QoS class scheduler; and

a IP packet transmitting section which transmits said IP packet to the lower layer, such as data link layer and network access layer.

[Claim 10]

A communication apparatus according to claims 8 and 9, wherein said input interface determines a class identifier in said IP packet based on a header data of the received said IP packet (a combination of the data of the layer 3 typified by the IP header, and to the data of the upper layer, layer 4, typified by TCP/UDP headers), and allocates IP-QoS (Internet Protocol Quality of Service) code.

[Detailed Description of the Invention]

[0001]

[Technical Field to which the Invention Belongs]

The present invention relates to a multi-layer class identifying communication apparatus and a communication apparatus used for IP network of a network layer of an OSI reference model.

[0002]

[Conventional Technique]

Recently, the Internet has become a de facto standard for international network and TCP/IP (Transmission Control Protocol/ Internet Protocol) is used as a base. Applying to the seven layers of the OSI reference model, IP functions as a network layer and TCP functions as a transport layer, and data are acquired by they are transferred from the Ethernet or a token ring as the lowest layer to IP, and then from TCP to an application layer.

[0003]

A router is located between a gateway and a repeater or a bridge as an intermediary device between LANs, storing frames received from a communication network, and acting as an intermediary device that forwards the frame to an appropriate communication network according to a destination network address in the frame data.

[0004]

[Problems the Invention Tries to Solve]

Basically, conventional routers do not perform priority control by packet, and all IP packet is processed equally. Although an IP packet includes an IP address in the header and is stored in a buffer of a router, read out from a buffer is controlled in FIFO (First In First Out) system and delay priority control scheme is not performed.

[0005]

Congestion occurs when IP packet concentrates to certain output port, leading to discard of packets in a buffer in a router. Generally, discard priority control to which packet to discard.

[0006]

A known ATM technique uses a concept of connection in which a route connecting a source address and a destination address is clearly defined by VPI/VCI (a virtual path identifier/a virtual channel identifier). This concept defines the quality of service (QoS) such as delay characteristics and discard ratios, which is necessary for a connection unit, and network apparatuses perform priority controls so that the QoS of connections are satisfied.

[0007]

(IP-QoS)

Recently, consideration is given to technologies that easily perform priority control on the Internet. Major examples are Intserve/RSVP (Resource Reservation Setup: A protocol enabling a network band control) and Diff-Serv (Differentiated Service). Diff-Serv performs priority control based on packet data alone as much as possible, while Intserve/RSVP simulates the concept of connection as introduced in the ATM.

[0008]

Not being applicable to a large-scale backbone network at low cost (i.e. the system lacking scalability), systems like Intserve/RSVP is not widely used. In order to solve the problem, the Diff-Serv makes primary consideration in scalability, availability at low cost, and adaptability to the high-speed performance of an OC-48 class (optical carrier) as an interface

for a backbone network optical fiber according to transmission rates.

[0009]

Instead of guaranteeing quantitative service as in the ATM-QoS, Diff-Serv sets a relative quality difference to facilitate the differentiation of service from the viewpoint of the Best-Effort. This is accepted as a substantially practical solution judging from specification progress by IETF (the Internet Engineering Task Force) as a Net-problem solving organization and vendor responses.

[0010]

(Diff-Serv)

The Diff-Serv has been discussed in the IETF, which regulates the Internet technologies. The Differentiated Service is a system for differentiating service levels, not for guaranteeing the QoS. It is basically a framework for (relative) priority control. That is, the Diff-Serv only defines the frameworks of QoS classes, and the details of the QoS classes and scheduling formats between the QoS classes are up to vendors and users.

[0011]

(Service Classes in Diff-Serv)

The Diff-Serv provides three types of traffics (service classes) including (1) Expedited Forwarding (Premium Service) (EF class), (2) Assured Forwarding Service (AF: Assured Forwarding class), and (3) Best Effort Service (BE: Best Effort class).

[0012]

The Expedited forwarding (EF: forward pressure, premium) class provides a virtual dedicated-line service such as an IP-CBR (Constant Bit Rate) on an IP network. Thus, it is necessary to perform precise transmission control including UPC (Usage Parameter Control) by additionally using a shaping section. Since the EF class is regarded as a class for a band guarantee service, it takes first priority over the assured forwarding service (AF class) and the best effort service (BE class), which will be described below.

[0013]

Unlike Expedited Forwarding (Premium Service) (EF) class, Assured Forwarding Service (AF) class is a framework of the (relative) priority control in it's own terms. The Assured Forwarding Service has four types of delay classes and three types of discard classes. The delay priority control is achieved by prioritizing order of packet transfer. It has a benefit of reducing transfer delay of applications sensitive to delay. The discard control is achieved by prioritizing discard of packets at

the congestion point of the network device.

[0014]

(3) Best Effort Service (BE class) corresponds to traffic besides EF class and AF class, and the priority control given to the BE class is the lowest in the service.

[0015]

Described above is the outline of the Diff-Serv. However, the recommendation of the Diff-Serv is constantly changing. Thus, the definitions and usage described above can be changed.

[0016]

Therefore, an object of the present invention is to provide a multi-layer class identifying communication apparatus flexibly adaptable to the function of a router connecting local-area networks by changing only parameters in accordance to service classes in a network layer of an ATM network handling IP packets.

[0017]

[Means for Solving the Problems]

Intending to solve the problem, the present invention is an apparatus that controls IP (Internet Protocol) packets according to the quality of the network service (Quality of Service: QoS), and realizes features described below.

[0018]

(1) In the input interface, class identifier in the device is determined according to the header data of the received IP packets, in particular, a combination of the data of the layer 3 typified by the IP header, and to the data of the layer 4, typified by TCP/UDP headers. A service quality code (Internet Protocol Quality of Service: IP-QoS) of each IP packet flow is allocated. To classify into 16 types; BE class, AF class as 4 types of delay classes and 3 types of discard classes and EF class, IP-QoS code on priority and scheduling is allocated by accessing SRAM from CAM searching.

[0019]

(2) To give versatility to IP-QoS code allocation (mapping of traffic classes), priority traffic is defined by classifying in any given combination of data in IP header and TCP header by priority and filtering as shown in Fig. 5 which will be later described.

[0020]

(3) As a scheduling method on transmission side, Weighted Round Robin Scheduling (WRRS) can be combined with a fixed priority scheduling method. Each IP-QoS class can be selected for the fixed priority scheduling, and a minimum frequency band can also be designated.

[0021]

(4) IP-QoS code determined on the reception side is stored in fixed length packet storage cell (Hereinafter referred to as an "Object" for convenience) of the apparatus and processed in the IP-QoS class determining section separate from packet (main signal) data.

[0022]

(5) Predetermined priority control is carried out based on IP-QoS code stored in the above-mentioned Object by performing in conjunction, so that an input interface and an output interface provided for input and output of a crossbar switch that switches to a destination address in the transmission-side perform scheduling according to the class.

[0023]

(6) Class queue management that aggregates and shares a plurality of IP flows is performed by processing by each set of IP packets classified based on IP-QoS code, and identical classes are processed in a same manner to obtain statistical multiplexing effect that allows a rapid processing.

[0024]

(7) Excessive traffic is restricted by monitoring traffics in units of the IP-QoS code at the input interface and the output interface, and deciding whether the IP-QoS code can be discarded or not.

[0025]

Also, in a communication apparatus including an input interface and an output interface in which a plurality of level of communication qualities can be selected, the input interface may include an IP packet receiving section which extracts the header data and TCP header data of the inputted IP packet; an IP-QoS class determining section which obtains class identifier to be determined by accessing to the IP-QoS code (class identifier) determination memory, using the header data of the IP packet header data as a search key; a reception-side switch interface control section which carries out a priority control that a destination has been specified based on a IP packet data received from the IP packet receiving section and on the IP-QoS class code corresponding to the IP packet determined in the IP-QoS class determining section; and a reception-side switch interface which issues a priority control function and a transmission request to the output interface card in units of the class identifiers. The IP-QoS class determining section monitors a coming traffic which exceeds a transmission permissive capacity set for every IP-QoS class, and carries out a discarding operation of IP packets or a

policing operation to lower transmission priorities of the IP packets when the coming traffic exceeds the transmission permissive capacity.

[0026]

Also, in a communication apparatus including an input interface and an output interface in which a plurality of level of communication qualities can be selected, the output interface which receives priority-controlled IP packets transmitted from the input interface may include a transmission-side switch interface which stores the IP packet received from the input interface in the transmission-side payload memory, and simultaneously generates IP packet data to write in the FIFO memory; IP-QoS class scheduler which carries out a queuing operation based on IP-QoS class code based on IP-QoS class code in the IP packet to primarily issue a transmission request, and control operation providing requested service quality by scheduling function based on WRR (Weighted Round Robin) method; a transmission-side switch interface control section which transmits the IP packet transmitted from the transmission-side switch interface according to the order scheduled by the priority of the IP-QoS class scheduler; and a IP packet transmitting section which transmits the IP packet to a data link layer and a network access layer.

[0027]

Also, in the above-mentioned communication apparatus, the class identifier in the IP packet is determined based on the header data of the received IP packet (a combination of the data of the layer 3 typified by the IP header and the data of the layer 4 typified by TCP/UDP headers representing higher-level layer), and the IP-QoS code is allocated to each of the IP packet flows.

[0028]

[Embodiments of the Invention]

The embodiments of the present invention will be described in detail referring to figures.

[0029]

[The First Embodiment]

(1) Description of the configuration

Fig. 1 is a block diagram of a multi-layer class identifying communication apparatus according to the present invention. An apparatus according to the present invention mainly includes input and output interface cards, an N×N crossbar switch 1g, and a switch scheduler 1h.

[0030]

An input interface line card includes an IP packet receiving section 1a, a reception-side switch interface control

section 1b, a reception-side switch interface 1d, a reception-side payload memory 1c, an IP-QoS class determining section 1e, and an IP-QoS code determination memory 1f.

[0031]

An output interface line card includes a transmission-side switch interface 1j, a transmission-side switch interface control section 1k, an IP-QoS class scheduler 1m, a class packet data queuing memory 1p, and an IP packet transmitting section 1q.

[0032]

Here, use of the input interface line card and output interface line card means that blocks equipped with required members are attached in the form of card to the crossbar switch 1g switching from a source address to a destination address, and these interface line cards may be simply provided as an input interface and an output interface. The multi-layer class identifying communication apparatus mainly includes the input/output interfaces and the switch scheduler 1h in addition to the crossbar switch 1g. In addition, in case of the communication apparatus with a simpler structure including neither the crossbar switch 1g nor the switch scheduler 1h that controls the switching of the crossbar switch 1g, the functions and software programs according to the embodiments of the present invention may be provided with a router handling IP packets on the Internet and supporting layers up to network layer, and a bridge having the function of the physical layer and a data link layer and carrying out filtering of the IP packets.

[0033]

(Input interface line card)

In an input interface line card (reception-side), IP packet receiving section 1a extracts an IP packet header data and a TCP/UDP header data accommodated in the higher-level layer, from data of IP packets transmitted after dividing packets of the application layer by TCP of the transport layer as the layer 4 of the OSI reference model.

[0034]

The IP packet receiving section 1a searches the IP packet for various kinds of conditions defined based on the contents of a CAM or SRAM area based on the received IP packet data, and the IP packet receiving section 1a retrieves processes such as (i) a queue priority, (ii) a filtering process (discarding/passing process), (iii) a SW priority/non-priority control process, and (iv) addition of a DSCP value of Diff-Serv from the IP packet as actions satisfying the conditions.

[0035]

The IP-QoS class determining section 1e accesses to the IP-QoS code (class identifier) determination memory 1f to acquire a class identifier, by using as a retrieval key the header data of the received IP packet (a combination of the layer 3 data typified by the IP header and the layer 4 data typified by TCP/UDP header corresponding to a higher-level layer). The IP-QoS class determining section 1e notifies the class identifier of the IP-QoS code to the reception-side switch interface control section 1b.

[0036]

Also, the IP-QoS class determining section 1e has means for monitoring an excessive traffic which exceeds a transmission permissive capacity set for every IP-QoS class, and for carrying out a discarding operation of packets (with IP-QoS code) or a policing (monitoring) operation to lower transmission priorities of the IP packets when the coming traffic exceeds the transmission permissive capacity, and decides a transmission frequency of the transmission packets according to an amount of network resource allocated to each IP-QoS class code that the packet belongs to.

[0037]

The reception-side switch interface section 1b carries out a priority control to the crossbar switch 1g based on packet data received from the IP packet receiving section 1a and by the IP-QoS code corresponding to the packet decided in the IP-QoS class determining section 1e.

[0038]

The reception-side switch interface control section 1d controls a transmission request in units of classes and of output interface cards, and always sends the transmission request of a higher priority to the switch scheduler 1h. This priority control is positioned as a delay priority control, in correspondence to a packet read operation from the reception-side payload memory 1c to crossbar switch 1g, and the IP packet is outputted to the crossbar switch 1g that selects and connects to the destination address.

[0039]

The receiving-side payload data memory 1c stores inputted IP packet data and empty packet data required when the IP packet is transmitted.

[0040]

(Output Interface Line Card)

The transmission-side switch interface 1j stores the packet data received from the crossbar switch 1g in the transmission-side payload memory 1i and simultaneously generates a packet data corresponding to the stored packet data to write in an FIFO memory 1n. The packet data is a virtual processing unit

defined in units of packets in the apparatus, and is hereinafter referred to as an "object". The object is defined to prevent the packet data from being carried around in the apparatus and is not the packet data itself. Packet transmission is carried out by passing the object in the apparatus.

[0041]

The FIFO memory 1n transmits a packet header data from the transmission-side switch interface 1j in a first-in first-out manner, keeping a predetermined delay time.

[0042]

The IP-QoS class scheduler 1m carries out queuing for every class based on the IP-QoS class codes contained in the objects. The IP-QoS class codes correspond with a plurality of delay classes and a plurality of discard classes, and the object stored in a queue having a high delay priority is primarily transmitted to the transmission-side switch interface control section 1k. The IP-QoS class scheduler 1m has a scheduling function based on the WRR (Weighted Round Robin) system to control in such a manner that a required service quality can be provided. In addition, on the IP network, precise transmission control including UPC (usage parameter control) and a Shaper is carried out to a premium service class providing a virtual dedicated line.

[0043]

The transmission-side switch interface control section 1k outputs the IP packets from the transmission-side switch interface 1j to the IP packet transmitting section 1q in the order of scheduling based on the priority determined by the IP-QoS class scheduler 1m.

[0044]

The IP packet transmitting section 1q outputs the IP packets to the lower layers such as the data link layer and the network layer of the Ethernet and token-ring networks.

[0045]

(2) Description of operation

Fig. 2 illustrates the functions of the apparatus according to the present invention. An example of the priority control applied in the present invention will be described using the configuration of the communication apparatus shown in Fig. 1.

[0046]

First, the IP-QoS class determining section 1e accesses to the CAM/SRAM to determine IP-QoS class (QoS codes stored in the apparatus) by using a predetermined data in the layer 3 or the layer 4 as a retrieval key. In this case, the IP-QoS class determining section 1e supports production of both Behavior

Aggregate (BA) class/Multi-Field (MF) class.

[0047]

Regarding the addition of EF class and policing control, a policing function for decided IP-QoS class (EF/AF1~AF4/BE) is supported. The policing control is basically the comparison between a token length of each class and the length of a transmitted packet. When the token length is shorter than the data packet length, the packet is discarded.

[0048]

The IP-QoS class determining section 1e transmits IP-QoS codes (class identifiers) to the reception-side switch interface control section 1b. The IP-QoS code determination memory 1f divides the IP-QoS codes into, for example, 16 kinds (5 × 3 + 1) of (EF. H), (AF1 to AF 4. H/M/L), and (BE. H/M/L) based on a delay class and a discard class. The IP-QoS code (class identifier) acquires the address of an empty area between the reception-side switch interface control section 1b and the reception-side switch interface 1d.

[0049]

The IP-QoS class scheduler 1m functions as a IP-QoS class scheduler, and carries out a process of outputting the object to the transmission-side switch interface control 1k from a high priority class EF > (AF1 to AF4/BE), for scheduling based on the IP-QoS class (in-apparatus QoS code) determined on the reception side. It should be noted that (AF1 to AF4/BE) carries out the scheduling based on the WRR system.

[0050]

Regarding the discard control in the IP-QoS class scheduler 1m, in case of AF1 to AF4/BE, the scheduler 1m compares the thresholds of the three classes of H/M/L with a buffer length to generate a drop object, and in case of EF, the IP-QoS class scheduler 1m supports by adding an H class.

[0051]

In terms of additions to the EF class and a shaping function, the shaping (delay scheduling) function is carried out to a determined IP-QoS class (only the EF class). The shaping control is basically a token-bucket system equivalent to the policing control. The token length of each class is compared with the length of a packet to be transmitted, and if the token length is shorter than the packet length, the transmission of the packet is postponed. The IP-QoS class scheduler 1m uses the class object queuing memory SRAM 1p to carry out the scheduling of objects.

[0052]

(Priority control)

As the switching priority control of the apparatus, the following four processes are assumed.

[0053]

(1) Mapping into an in-apparatus delay class by the reception-side switch interface control section 1b

The mapping into the in-apparatus delay class is a method for transmitting a transmission request. In this method, designation of delay priority control for the two classes (H/L) corresponds to six delay classes.

[0054]

Before transmitting a packet to an output IF card (line card on the transmission-side), the reception-side switch interface 1d issues a connection request to the switch scheduler 1h. The switch scheduler 1h arbitrates transmission requests from input line cards and notifies the connection data of input and output paths to the crossbar switch 1g. In addition, the scheduler 1h issues a result notification of connection arbitration to each of the input line cards on the reception side.

[0055]

The reception-side switch interface 1d controls the requests of each class and each output IF card. The reception-side switching interface 1d always transmits a high priority request to the switch scheduler 1h primarily. This priority control is positioned as the delay priority control, corresponding to the packet read operation from the reception-side payload memory 1c to crossbar switch 1g.

[0056]

In case of mapping into the in-apparatus delay class, the reception-side switch interface control 1b positioned on a previous stage of the reception-side switch interface 1d determines the delay priority according to the traffic classes of packets. In Fig. 2, in order to match with the number of classes compliant with the Diff-Serv, six kinds of delay priority classes are determined as the traffic classes of packets. An example of matching of the two kinds (High/Low) of class queues with respect to the crossbar switch 1g is shown.

[0057]

(2) Mapping into a discard class carried out when an input packet is written in the reception-side payload data memory 1c

In the acquiring process of the address of an empty area (a free page address) of mapping into the discard class, the discard priority control of three kinds of classes (H/M/L) are designated. In this manner, the three kinds of discard classes

correspond only with the EF class.

[0058]

When the input packet is written in the reception-side payload data memory 1c, the reception-side switch interface control section 1b positioned in the previous stage acquires the address of the empty area (free page address) in the reception-side payload data memory 1c. With the use of the free page address, the input packet is written in the reception-side payload data memory 1c.

[0059]

Concerning the acquisition of the empty-area address, the reception-side switch interface 1d has several kinds of priorities. To describe simply, Fig. 2 shows an example in which three kinds of discard classes (H/M/L) are used.

[0060]

The reception-side switch interface 1d monitors the capacity of the empty area of the reception-side payload data memory 1c. When the capacity is smaller than a predetermined high threshold, the reception-side switch interface 1d permits only the writing of a high priority packet. When the capacity is smaller than a predetermined low threshold, the reception-side switch interface 1d permits only the writings of an intermediate priority packet. Except for these cases, the reception-side switch interface 1d permits the packets of any classes to be written in the memory. This control is the discard control to the memory 1c, and corresponds to the three kinds of discard classes of the traffic classes.

[0061]

(3) The priority control carried out when an output packet is read out from the transmission-side payload memory 1i (delay priority control)

The packet read operation from the transmission-side payload data memory 1i is controlled by the IP-QoS class scheduler 1m, the transmission-side switch interface control 1k, and the transmission-side switch interface 1j.

[0062]

The priority control of the communication apparatus is a control of an order in which packets are read out from the transmission-side payload data memory 1i, and is comparable to delay priority control. The delay priority control handles six kinds of delay classes of IP-QoS codes.

[0063]

(4) Priority control carried out when a packet transmitted from the crossbar switch is written in the

transmission-side payload data memory (discard priority control)

The packet transmitted from the crossbar switch 1g is written in the transmission-side payload memory. The data of the written packet is notified as objects to the IP-QoS class scheduler 1m. The IP-QoS class scheduler 1m controls queue lengths in the transmission-side payload data memory in units of classes, and compares the queue length with a threshold of the discard class to determine whether packets contained in the transmission-side payload memory should be discarded or not.

[0064]

As a result, this control carries out a discard priority control. The discard priority control corresponds to the discard class of the traffic class.

[0065]

The transmission-side switch interface control section 1k carries out packet transmission/packet discard by using two kinds (transmission/discarding) of objects. In the transmission of a packet, a packet staying in an FSU memory is read out and transmitted in response to an issued read command. In case of discarding a packet, a packet staying in the FSU memory is discarded in response to an issued drop command.

[0066]

(IP-QoS Class Determining section)

Fig. 3 illustrates the main part of the IP-QoS class determining section 1e. The details of the IP-QoS class determining section will be described with reference to Fig. 3.

[0067]

As shown in the figure, the IP-QoS class determining section 1e is composed of a header extracting section 3a, a header checking section 3b, an IP-QoS code search section 3c, a policing control section 3d, an IP-QoS code output section 3e, and a parameter register control section 3f.

[0068]

The header extracting section 3a extracts a predetermined data based on the formats of the IP header and TCP/UDP header of IPv4 shown in Fig. 9, and transmits fields included in the extracted data as IP data to the IP-QoS code search section 3c.

[0069]

In Fig. 9,

(1) a densely hatched section (Ver) shows a field to be checked.

[0070]

(2) Roughly hatched sections (TOS, Src IP Address Dst, IP Address, L4 Src Port, and L4 Dst Port) show fields for specifying

classes as objects of a search key.

[0071]

The extracted data includes a 4-bit version (Ver), an 8-bit TOS (type of service) identifier, a source (Src) IP address, a destination (Dst) IP address, a L4 Src port number of a layer 4 header, and a L4 Dst port number of the layer 4 header.

[0072]

An Internet header length (IHL) which indicates the size of the IP header, a datagram length which indicates the total length of the entire packet including the IP header and IP data, and an identification which indicates an identifier restoring a fragment are included. Also, a Flag M is composed of 3 bits, a 13-bit fragment offset indicates the location of a fragment after division in original data, and a time to live indicates a time during which the presence on a network is permitted. In addition, a protocol specifies the upper layer protocol, and a header checksum indicates the check sum of the IP header.

[0073]

The IP header checking section 3b checks the normality of the IP header and outputs the result of an IP Header Error or Encap Field to the IP-QoS code output section 3e. The IP-QoS code search section 3c accesses a CAM (Contents Addressable Memory) and SRAM by using data received from the header extracting section 3a as a search key to determine an IP-QoS code. The determined IP-QoS code includes data such as a class identifier of the apparatus and priority in switching control.

[0074]

The policing control section 3d monitors the traffic of each class determined by the IP-QoS code search section 3c positioned at the previous stage to control or restrict an excessive traffic flow. In this processing, the token-bucket system is used to monitor traffic violation. In the token-bucket system, a token quantity contained in a bucket increases with a ratio calculated based on an expression: T (elapsed time) × r (average rate). On arrival of a packet, if the present tokens do not have a length long enough to contain the received packet, the packet is discarded.

[0075]

The above processing is based on a simple logic. When the token length is shorter than the packet length, the value of the discard bit is set to "ON" to indicate the packet to be discarded, and the packet is transmitted to a rear-stage block. In the case except for the above, or of the token length is the same as or longer than the packet length, the discard bit is set

to "through".

[0076]

In order to quickly determine whether the traffic violation is caused or not, the following method may be employed when producing and adjusting hardware and software as an implementation.

[0077]

The condition for passing an input packet is set to "token is equal to or larger than 0" instead of "token is equal to or larger than packet length".

[0078]

After transmission of the packet, the token quantity may be negative since the quantity corresponding with the length of the packet is subtracted from the present quantity of tokens. When it is a negative value, the transmitted packet is regarded as an object causing traffic violation. The use of the determination circuit simplifies the configuration by determining violation based on token code data (1-bit data).

[0079]

The IP-QoS code output section 3e carries out the re-timing of an IP-QoS code determined by the IP-QoS code search section 3c, a filtering bit from the policing control section 3d, and error data from the header checking section 3b to output to the rear-stage block (reception-side switch interface control section 1b). An operator determines how to combine fields included in the mapping (MF/BA classifier) IP headers of traffic classes and how to make the combined fields correspond with the traffic classes. These issues are not specified in the recommendation of RFC of the IFTF. In order to provide the above correspondence with flexibility, it is necessary to carry out the mapping of traffic classes based on arbitrary combination of extracted header data.

[0080]

For example, when the priority control is carried out to the IP traffic specified between certain contract users, the determination of traffic classes is carried out based on the combination of a source (Src) IP address and a destination (Dst) IP address. In the specified IP traffic, when the priority control is carried out only to the traffic of HTTP (the protocol for exchanging hypertexts with a WWW server on the Internet), it is necessary to classify the traffic based on a combination of a Src Port number and a Dest Port number contained in the header of the upper layer. In addition, when the priority control is carried out only to the traffic from a certain server, it is

necessary to classify the traffic by referring only to an Src IP address as the IP address of the server and an Src Port number. As shown here, when the traffic is classified based on a combination of the plurality of fields of the IP header and the upper layer, this method is called a multi-field (MF) classifier.

[0081]

Other than the above method, there is a classifying method called a behavior aggregate (BA) classifier. The BA classifier determines one of the traffic classes by referring only to the TOS field of the IP header. The TOS field defined in the IP header is a special field defined for the Differentiated Service. The TOS field is used to reduce a procedure for determining the traffic based on a combination of fields of the IP header in a router. That is, an upper-stage router determines the traffic classes based on the fields of the IP headers and adds the class data to the TOS field to be transmitted. Next-stage routers subsequent to the upper-stage router only need to carry out the priority control to each traffic class by referring to only the TOS fields. However, vendors need to determine how to use the TOS fields. Thus, restrictions may involve such as the necessity to connect the apparatuses of the same affiliates (vendors) to each other, and the need to make an operational rules between adjacent routers.

[0082]

As shown above, mapping of the traffic class requires a method that enables registration according to the various kinds of combinations of arbitrary fields. In the present invention, the MF classifier and the BA classifier can be simultaneously supported by the following method.

[0083]

(Operation of the Class Search Section)

The operation of the class search section will be described with reference to Figs. 4 and 5.

Fig. 4 shows the structure and operation of the class search section 3c. Fig. 5 shows a searching operation flow. As shown in Fig. 4, the IP-QoS class search section 3c is composed of a layer L4 port converting SRAM, a source IP address search CAM, a destination IP address search CAM, a priority mapping search CAM, and a priority setting SRAM. IP-QoS class search section 3c extracts a software priority control data. In addition, the flow chart shown in Fig. 5 includes a CAM search section inputting the source IP address, the destination IP address, a TOS identifier, and a protocol, to output a searched address Q, and a SRAM access section inputting the searched address Q, the upper layer TCP

source port, and the upper layer TCP destination port to output an IP-QoS code.

[0084]

In this situation, mainly the condition data is described on the CAM, and the action data is described on the SRAM. The actions described on the SRAM correspond with the conditions described on the CAM. Thus, dividing both CAM and SRAM as recording mediums is not necessarily a requirement, but considering the high-speed operation of the SRAM, the CAM may be formed by dividing the memory area of the SRAM. In this case, it is primarily necessary to apply an implementation method for effectively using the resource (area: the number of entries) of the CAM.

[0085]

As described above, the MF classifier classifies traffic classes based on the arbitrary combinations of an Src IP address, a Dst IP address, an Src Port number, a Dst Port number, a protocol number, and a TOS. Regarding the MF classifier, the simplest searching method is a method in which the values of a Src IP address, a Dst IP address, a Src Port number, a Dst Port number, a protocol number, and a TOS field are set as registered data on the CAM capable of designating a mask for each entry, and the searching operation is carried out based on the packet header data every time a packet is input. However, since there are restrictions on the bit width of the CAM, practical ideas for implementation are needed.

[0086]

The present invention has a sequence for acquiring an IP-QoS code by using the general-purpose CAM and the following two-stage searching method. The BA classifier carries out the process of converting to an IP-QoS code by referring to only TOS fields in the same framework as the process carried out by the MF classifier. That is, the process by the BA classifier can be regarded as a case where only the TOS fields are usable in case of the MF classifier. Considering the configuration of the CAM, it is possible to operate using both classifiers.

[0087]

(1) First, as a prior processing, fields are degenerated into key values to be registered in the CAM and the SRAM.

[0088]

(2) Next, as a classification process, entries as the registered key values are searched on the CAM.

[0089]

A detailed explanation will be given below with reference

to Figs. 4 and 5.

[0090]

(Prior Registration in the CAM)

Prior Processing 1: the registration/degeneration of a Src IP address (shown in ① of Fig. 4)

All Src IP addresses/prefixes included in the entries for classification are registered in advance. In the prior process, when the Src IP address and the Dst IP address are degenerated, the Src IP address is used as a search key to carry out the searching operation by use of the longest prefix match under regulations based on the classless inter domain routing (CIDR). A CAM address obtained through the searching operation is set as the address A. When no key value is registered, all "0" are a value indicating the address A. This process is equivalent to step (1) shown in Fig. 5.

[0091]

Prior Processing 2: the registration/degeneration of a Dst IP address shown in ② of Fig. 4

In the same way as the prior processing 1, all Src IP addresses/prefixes included in the entries for classification are registered in advance. The Dst address is used as a search key to carry out retrieval by the longest prefix match. A CAM address obtained by the searching operation is set as Addr_B. When no key value is registered, all "0" are the value of the Addr_B. This process is equivalent to the process of step (2) shown in Fig. 5.

[0092]

When the CAMs are connected in parallel to each other, steps S1 and S2 shown in Fig. 5 can be carried out in parallel.

[0093]

Prior Processing 3: the registration/degeneration of port numbers (the data of the layer 4 shown in ①′ and ②′ of Fig. 4.)

The purpose of the processing 3 is to determine the layer 4 application of from the Src port number or the Dst port number to degenerate into a predetermined key value. The port numbers are classified into well-known port numbers defining the protocol of the layer 4 and numbers arbitrarily added by terminals.

[0094]

In the normal layer 4 application, the well-known port number of the layer 4 application executed by a server is added to the Dst port number of a packet routed to a server from a client. The well-known port number is registered in a memory (table).

[0095]

Since the well-known port numbers required in the operation are limited to a few kinds of numbers (HTTP, TELNET, FTP,

etc.), only a small amount of memory (256 × 8 = 2064 bit) is needed.

[0096]

In the processing 3, the port number is used as the address and access is carried out to the memory for converting from the port number to a key value to acquire a predetermined key value. The key is composed of a code specifying the layer 4 and a flag designating whether the key value is valid for an Scr port or a Dst port. The reason for designating the flag for the Src/Dst ports is to enable the classification of a one-way traffic, such as treating the class of traffic to a server treated as a high priority class, while treating the class of traffic to a client as the best effort class.

[0097]

As shown in steps 3 and 4 of Fig. 5, the processing is implemented on each of the Src port number and the Dst port number. The reading of a conversion key based on the Src port number is equivalent to the step 3 of Fig. 5. The reading of a conversion key based on the Dst port number is equivalent to the ①′ and ②′ of Fig. 4 and the step 4 of Fig. 5.

[0098]

The results of both processes mentioned above generates a key value into which a well-known port number is degenerated and a flag showing whether the key value comes from an Src port or a Dst port. This process is equivalent to the step (5) of Fig. 5 for calculating the Pory Key.

[0099]

(Class determining process shown in ③ and ④ of Fig. 4)

The CAM is searched again based on key values, TOS fields, and protocol numbers obtained in the prior process shown in the steps (1) to (5) of Fig. 5. This process is equivalent to a step (6) of Fig. 5. The combination of keys, that is, keys to be used for the searching operation as valid data, are defined as mask data for each entry in the CAM. This process can provide a CAM address (addr_Q) in which each entry is stored.

[0100]

Lastly, in a step (7) of Fig. 5, the CAM address (addr_Q) is used as pointer data (address) for an external memory. As a result, ultimately required traffic class data can be obtained.

[0101]

The reception-side switch interface control section 1b maps data (IP_ INFO) obtained in the operation flow shown in Fig. 5 into an object and transmits to the reception-side switch interface 1d. Then, the input IF cards carry out a variety of

priority control.

[0102]

It should be noted that when there is no hit in the registered data through the searching operation of the CAM, the CAM returns an address "0". This is defined as a class code indicating BE traffic. In the external memory, the value of the external memory at a reset time is set to be all "0", and the area of the address "0" stores the data of BE traffic class. Thus, traffic which is not hit in the CAM searching operation is regarded as BE traffic. The details of tables stored in the CAM are shown in Figs. 6A to 8.

[0103]

Fig. 6 shows a configuration example of a table stored in the memory CAM (Contents Addressable Memory) and a table of a storage area for IP Src Prefix entries. Fig. 7 shows a table of a storage area for IP Des Prefix entries and a table of a storage area for IPINFO address entries. In addition, Fig. 8 shows a table of a storage area for IPINFO entries.

[0104]

In the cases of Figs. 6 to 8,
D: Discard Indication (0: normal, 1: discard)
is used in filtering.

[0105]

In addition,
P: Packet Priority (0: low priority, 1: high priority);
Route: upper 1-bit route change flag (0: default route setting, 1: lower 4-bit route field setting);
Output TOS
bit 9: DSCP update flag (0: non-rewriting of the TOS DSCP Field bit 7-2, 1: implementation of rewriting thereof);
bit 8: CU update flag (0: non-rewriting of TOS CU Field bit 1 and 0, 1: implementation of rewriting thereof);
bit 7-2: TOS DSCP (differentiated service code point) field;
bit 1-0: TOS CU (currently unused) Field.

[0106]

(Traffic Regulating method)

As shown in Fig. 15, in the traffic regulating method of the present invention, traffic characteristics are expressed by a token-bucket model. The token-bucket model is applied in (1) the policing section 3d of the reception-side interface section and (2) the shaping section 8f of the transmission-side interface section.

[0107]

That is, algorism for detecting traffic violation has a

circuit structure equal to a shaping and policing section. Fig. 15 shows the token-bucket model. The token-bucket has a token-containing register (bucket b) to add tokens in the bucket cyclically, and tokens increase at an average rate (r). In the token adding process, the value of WC + token is compared with the token upper limit value b, and when the value exceeds the upper limit value b, the value b is set.

[0108]

As a condition for transmitting a packet, it is necessary that a token giving a size corresponding with the size of a packet to be transmitted is present in the bucket. After the packet is transmitted, the number of tokens giving a length equivalent to the length of the transmitted packet is reduced from the present number of tokens. On packet transmission, if a token having a length equal to or greater than the packet length is not present in the bucket, the transmitted packet is a traffic violation object. When the condition is applied to the policing section 3d, the packet as the traffic violation object is discarded, or, for example, low priority marking is carried out. In case of the shaping section 8f, transmission of the packet is waited until tokens are accumulated to permit the transmission of the traffic violation packet.

[0109]

(Additional Description shown in Fig. 16)

The traffic violation occurs when there are not enough tokens accumulated to transmit a packet in the bucket. The number of tokens contained in the bucket increases at a ratio calculated by an expression: T (elapsed time) × r (average rate). When tokens accumulated in the bucket do not reach a required length on receiving a packet, the transmission of the packet is postponed until shortfall of tokens are accumulated to required length.

[0110]

Fig. 16 shows the operational images of policing and shaping. With reference to Fig. 16, a description will be given as follows.

[0111]

(Example of Policing Operation)

(1) Assume that packet 1 arrives at time T1. Since this is not traffic violation pattern, the packet 1 is outputted at the same time T1'. The number of tokens left in a bucket at the time of output is indicated by X1.

[0112]

(2) Assume that packet 2 having size s2 arrives at time T2. The number of tokens necessary to output the packet 2 is

equivalent to S2. However, it is supposed that S2 > X1 + (T2 — T1) × r. In this case, a traffic violation is caused, and the packet 2 is discarded.

[0113]

(Example of Shaping Operation)

(1) Assume that packet 1 arrives at time T1. Since a traffic violation is not caused, the packet 1 is outputted at the same time T1'. The number of tokens left in the bucket at the time of output is indicated by X1.

[0114]

(2) Assume that packet 2 having size s2 arrives at time T2. The number of tokens necessary to output the packet 2 is equivalent to s2. However, it is supposed that S2 > {X1 + (T2 — T1) × r. In this case, a traffic violation is caused. Thus, the packet 2 is outputted after waiting for a time τ during which a required number of tokens are accumulated in the bucket. In this case, the time τ can be calculated by an equation s2 = X1 + {(T2 + τ) — T1} × r.

[0115]

(Details of Priority Control by Output Interface Cards)

With reference to Fig. 2, a description will be given of the priority control by the output interface (IF) cards. Regarding the output IF cards, an IP-QoS class scheduler 1m carries out predetermined priority control based on IP-QoS codes (traffic classes) defined in object data. The IP-QoS codes are defined as data for identifying traffic classes, and show a plurality of delay classes and a plurality of discard classes. The priority control carried out by the IP-QoS class scheduler 1m is equivalent to ③ and ④ of Fig. 4. Fig. 10 shows a concept view of the priority control by the IP-QoS class scheduler 1m.

[0116]

(Priority Control by the IP-QoS Class Scheduler)

The IP-QoS class scheduler 1m receives objects via a FIFO 1n from the transmission-side switch interface 1j and a classifier 8a classifies the objects. The classified objects are stored in a common buffer 8c selecting a queue for each class. The common buffer 8c stores the objects for each class. There are six kinds of object queues 8c in total: classes EF, AF1 to AF4, and BE. These object queues are all controlled by the common buffer 8c. The definitions of the EF, AF, and BE classes have already been provided in the description of the conventional art.

[0117]

Since the EF class is regarded as a band guarantee service, precise transmission control is carried out by

additionally using a shaper 8f. A Weighted Round Robin controller 8d shown in Fig. 10 carries out selector control to an output class selector 8e so that the EF class queue is primarily read before the AF/BE class queues are read. The shaping section 8f has a circuit structure composed of a token-bucket model equivalent to a policing unit. The difference between a shaper section and a policing section is only on whether the transmission of the packet having a traffic violation is postponed or the packet is discarded. As in the case of the policing section in terms of implementation, transmission process can be carried out at high speed when the packet transmission is determined by using only bits added to tokens. Furthermore, the circuit structure can be simplified.

[0118]

The scheduling of the AF and BE classes is carried out by the Weighted Round Robin (WRR) system.

[0119]

(Additional Explanation of WRR)

The WRR scheduling is an expanded round robin scheduling system. According to a predetermined weight ratio, service for each class is provided. In the WRR system, each class has a counter. Each counter shows the number of cells (or packets having fixed lengths) routable by the time the counter is reset. On resetting, the value of the counter is set to be the weight value of each class. When the counter value of a selected class is equal to or greater than "0" and the buffer includes one or more cells, one cell of the class is output to decrement the counter value. When the weight values of all of the classes are "0" or the number of cells in the buffer is "0", all of the weight counters are reset. Thus, when all classes have sufficient input traffics, the number of output cells corresponds with a weight ratio.

[0120]

Each of the AF class and the BE class has a weighted counter 8g and a weight value. A Weighted Round Robin controller 8d determines which objects of the AF/BE classes should be read based on the weight values and the data of the queue lengths of class queues 8c of a common buffer to control the selector. The queues of the selected AF class and BE class output the objects thereof in the order of FIFO.

[0121]

(Discard control)

A discard control logic unit 8b shown in Fig. 10 carries out discard control based on a logic shown in Fig. 11. The queue

length 8h of each class is controlled based on the input object data. A discard control threshold 8k is defined for each class, and the threshold of the entire common buffer 8c is also defined. The discard control logic 8b carries out discard control by comparing the queue length with the thresholds 8k and notifies an object which should be discarded as a drop object to a transmission-side switch interface 1j.

[0122]

In addition to the discard control carried out by the above comparison between class unit queue length 8h and threshold, another discard control is carried out by comparing the sum of the entire queue lengths of the same discard class with the thresholds thereof in the overall common memory 8c. In Fig. 11, three kinds of discard classes and three thresholds corresponding therewith are provided. Thus, (1) no discard control is carried out over all of the classes; (2) discard control is carried out to only a low priority class; (3) discard control is carried out over both a low priority class and an intermediate priority class; and (4) discard control is carried out over all of the classes. In Fig. 11, when the value of the class queue length becomes equal to or larger than the value of threshold 3, all packets are discarded, regardless of whether there is an empty area or not.

[0123]

In the same way, the threshold is defined not only in the AF-class unit but also with respect to the sum of discard classes in the entire common memories to carry out discard control.

[0124]

(Buffer used for the WRR system)

Fig. 14 illustrates the main part of the buffer used for the WRR system. With respect to class object queues 8c, weight counters (111, 121, ...1n1) controlling the present weight values and preliminary determination counters (112, 122, ... 1n2) controlling weight values after the weight counters are reset are provided.

[0125]

When class determination cannot immediately be made in the first processing by the Weighted Round Robin, the similar class determination processing is made with the weight value of the preliminary determination weight counter obtained after the reset of the weight counter. When this processing still cannot make the class determination, fixed delay priority control is carried out to select the AF classes. The above class determination processings are carried out in parallel while excluding a loop processing. Thus, the processings can be carried

out at high speed.

[0126]

(Flow of Scheduling Operation)

Each of Figs. 12 and 13 shows the flowchart of the scheduling. As processing paths, an EF class determination path and an AF/BE class determination path exist. Since the EF class has absolute precedence over the AF/BE classes, the EF class packet is selected when there is a transmission request for an EF-class packet. Only when there is no transmission request for an EF-class packet, a predetermined AF/BE class is selected.

[0127]

Regarding AF/BE class determination, the Weighted Round Robin system carries out first and second determinations, and fixed delay priority control is carried out to select from the AF/BE classes as a third determination.

[0128]

When the processing flows shown in Figs. 12 and 13 are read, the following points need to be considered. Each parameters in each of Figs. 12 and 13 are described as follows. In the EF class, W[EF] indicates an initial weight (added token value) of the EF class; WH[EF] indicates an upper limit of token value; WC[EF] indicates a token value (variable) of the EF class; and Add Time[EF] indicates a token addition cycle, which is, for example, set by an average rate under contract). In the AF/BE classes (i AF1, 2, 3, and 4, and BE), W[i] indicates a class initial weight value; WH[i] indicates a weight-counter upper limit value; WC[i] indicates a weight counter value (variable); WC[i]r indicates a weight counter value (variable) after a single reset of the weight counter; WC[i]r WC[i] + _ class ptr indicates a pointer showing a transmitted/searched class; MPSZ (Maximum Packet Size) applied in the AF/BE classes indicates a packet maximum length used as a judgment reference for rare cases; and Q[i] > 0 indicates the absence or presence of an object in a queue. In addition, parameter fixed priority (Q[i]) indicates a situation in which an output request is transmitted in fixed priority order by AF1, AF1, AF2, AF3, AF4, and BE only based on the parameter Q[i] > 0 regardless of a WC value. Parameter Length indicates the length of a packet and is used as a common parameter among all of the classes.

[0129]

First, the process of EF-class selection will be described below.

[0130]

(Cyclic Addition of Token)

In Fig. 12, as an initial processing, in step S0a, a token WC[EF] is initialized to be "0". In step S0b, a token addition timing is generated. When it is a predetermined timing (cnt = = Add_time [EF]), a token is added in step S0c. An initial token W[EF] is added to the present WC[EF]. Next, the token addition cycle is, for example, set by an average rate under contract. Then, the value of WC[EF] + W[EF] is compared with the token upper limit value (WH[EF]). When the value of WC[EF] + W[EF] exceeds the value WH[EF], the WH[EF] is set.

[0131]

(Generation of EF-class Sending Requests)

Next, as shown in Fig. 13, when an object reading request is received (step S1a), the queue length is determined in step S1b, and the presence or absence of a token is determined in step S1c.

[0132]

In the above determination process, when the queue of the EF class has one or more objects and one or more tokens are present, a sending request (Send_req [EF] = ON) is routed in step S1d. Next, after a packet is transmitted, tokens giving a length equivalent to a transmitted-packet length are reduced in step S1e. In the final-stage selector processing as step S3, a sending request for the EF class as the highest priority class is accepted.

[0133]

Next, an explanation will be given of the AF/BE class selection processing.

[0134]

(AF/BE Class Selection Operation)

The AF/BE class selection processing will be illustrated below with reference to Figs. 12 and 13.

[0135]

(1) The values of both class pointer and weight counter are initialized (steps S0d and S0e).

[0136]

(2) After an object reading request is received in S1a, only when the EF class object cannot be transmitted, that is, only when the results of the determinations obtained in steps S1b and S1d are both "No", the AF/BE class selection is carried out. In step S2b, the queue length is determined, and in step S2c, weight validity is determined.

[0137]

(3) Since there is provided a plurality of AF/BE classes, the above determination processing (2) is carried out to all of the classes in step S2e. In step S2g, a class pointer (a pointer indicating a searched class) is updated in accordance with the

order of Round Robin.

[0138]

(4) When there is a class corresponding with sending conditions :queue length > 0 and weight value > 0, a sending request of the corresponding class (Send_req [Class] = ON) is routed in step S2d.

[0139]

The above operations (1) to (4) are class selection by a first routine.

[0140]

(5) In the above retrieving processing (3), when all of the classes do not correspond with the sending conditions, an initial weight value defined for each class is added to the present weight value in all of the classes. This process is called a weight-counter reset processing.

(6) When class selection cannot be carried out in the first routine, that is, when there is no sending class candidate, class selection is carried out again by a weight value obtained after a single retransmitted of the weight counter reset in the processing (5). As shown in Fig. 11, there are a register (WC[class]) controlling a present weight value and, additionally, a register (WC[class]r) controlling a weight value after the single reset of the weight counter. As in the case of the processing (4), when there is a class corresponding with the sending conditions: queue length > 0 and weight value > 0, a sending request of the corresponding class (Send_req [Class] = ON) is output.

[0141]

The above operations (5) and (6) are class selection by a second routine.

[0142]

(7) In the second routine, when class selection cannot be carried out, that is, when there is no output class candidate, fixed priority control is carried out for class selection as an exceptional case. The purpose of the process is to avoid a situation in which no packet cannot be routed although there is an empty space in the capacity of a line. A class to be outputted is selected based on a predetermined fixed priority. The processing does not depend on a weight value. In this case, an output candidate is determined only by determining whether an object is present or not in a queue. With this processing, two or more weight counter resets never happen in a single class selection processing. A normal determination is carried out by the Weighted Round Robin system, while fixed delay priority (which is called a

fixed priority mode in Fig. 12) is given to carry out an exceptional determination.

[0143]

(8) An output class is selected by a sending request for the corresponding class selected in the operations (4), (6), and (7) in step S2n.

[0144]

In the operation of step S2n, the maximum packet size MPSZ is used as a parameter. This is because it is necessary to avoid a situation in which when a packet having an excessive length arrives, the value of the weight counter becomes a large negative value. For example, the maximum length of the IP packet is 64 KB, which can be unusual as the size of a transmitted packet. Statistically, a usual packet length is assumed to be a few kilobyte at largest. Thus, arrangement is made to enable to set a maximum packet length regarded as an unusual packet length. In this arrangement, when a very large packet arrives, the arrived packet length is compared with the maximum packet length. When the arrived packet length is larger than the maximum packet length, it is regarded as an exceptional case. As a result, the reduction processing of a weight counter is skipped, or the weight value is forced to be "0" to impose a certain penalty.

[0145]

According to the above embodiment, in the Diff-Serv whose specifications are still in flux, the TCP layer of the OSI reference model is compared with the IP layer thereof. From IP packets, IP-QoS codes are allocated independently from the IP packet. Preferably, various kinds of communication service qualities are classified based on the IP-QoS codes. The classification of the IP-QoS codes permits traffic congestion in communication systems to be relieved.

[0146]

[Effect of the invention]

According to the present invention, the following advantages are attained.

[0147]

(1) An operator can set prioritized traffic by combining the packet data of the layer 3 and that of the layer 4 (the flow unit of each upper application).

[0148]

(2) Assuming versatile operation, Weighted Round Robin Scheduling (WRRS) can be combined with a fixed priority scheduling system. Each QoS class can be selected by the fixed priority scheduling, and a minimum frequency band can also be designated.

[Brief Description of the Drawings]

[Fig. 1] A schematic diagram of a multi-layer class identifying communication apparatus according to the present invention.

[Fig. 2] A block diagram showing a structure of a multi-layer class identifying communication apparatus according to the present invention.

[Fig. 3] A diagram showing an operation of the multi-layer class identifying communication apparatus according to the present invention.

[Fig. 4] A diagram showing a main portion of an IP-QoS class determining section according to the present invention.

[Fig. 5] A flow chart showing a process of the class search section according to the present invention.

[Fig. 6] A diagram showing a table structure of a CAM region division according to the present invention and a diagram showing the structure of an IP Src Prefix entry storage region.

[Fig. 7] A diagram showing a table block structure of the IP Src Prefix entry storage region according to the present invention, and a diagram showing the structure of the IP Src Predix entry storage region.

[Fig. 8] A diagram showing a table block structure of an IPINFO entry storage region according to the present invention.

[Fig. 9] A diagram showing a format of IPv4 & TCP/UDP/Other Header used in the present invention.

[Fig. 10] A principle diagram of an IP-QoS class scheduler according to the present invention.

[Fig. 11] A diagram showing a discarding control logic according to the present invention.

[Fig. 12] A flow chart showing a process of the scheduler according to the present invention

[Fig. 13] A flow chart showing a process of the scheduler according to the present invention

[Fig. 14] A diagram showing a main portion of a WRR object buffer according to the present invention.

[Fig. 15] A diagram showing policing and shaping in a token-bucket model according to the present invention.

[Fig. 16] A diagram of the policing and shaping operation according to the present invention.

[Description of the Reference Numerals]

1a IP packet receiving section

1b reception-side switch interface section

1c reception-side payload data memory

1d reception-side switch interface control section

1e IP-QoS class determining section
1f IP-QoS code (class identifier) determination memory
1g crossbar switch
1h the switch scheduler
1i transmission-side payload data memory
1j transmission-side switch interface section
1k transmission-side switch interface control section
1m IP-QoS class scheduler
1n FIFO
1p SRAM
1q IP packet transmitting section
3a header extracting section
3b header checking section
3c IP-QoS code search section
3d policing control section
3e IP-QoS code output section
3f parameter register control section
8a classifier
8b discard control logic unit
8c common buffer
8d Weighted Round Robin controller
8e output class selector
8f shaping section
8g weight counter
8h class queue length
8j total queue length
8k threshold (class, level, total)

[Document Name] Abstract

[Abstract]

[Problems] Provides a multi-layer class identifying communication apparatus flexibly adaptable by changing parameters.

[Solving Means] An input interface of the apparatus determines class identifier in the apparatus from the header data (a combination of the data of the layer 3 typified by the IP header, and to the data of the upper layer, layer 4, typified by TCP/UDP headers)of the received IP packet, and allocating IP-QoS (Internet Protocol Quality of Service) code dealing each IP packet flow in the apparatus, and with respect to the above-mentioned IP-QoS code allocation (mapping of traffic class), priority traffic is defined based on arbitrary combinations of a plurality of fields in IP header and TCP header.

[Selected Drawing] Fig. 1

Fig. 1

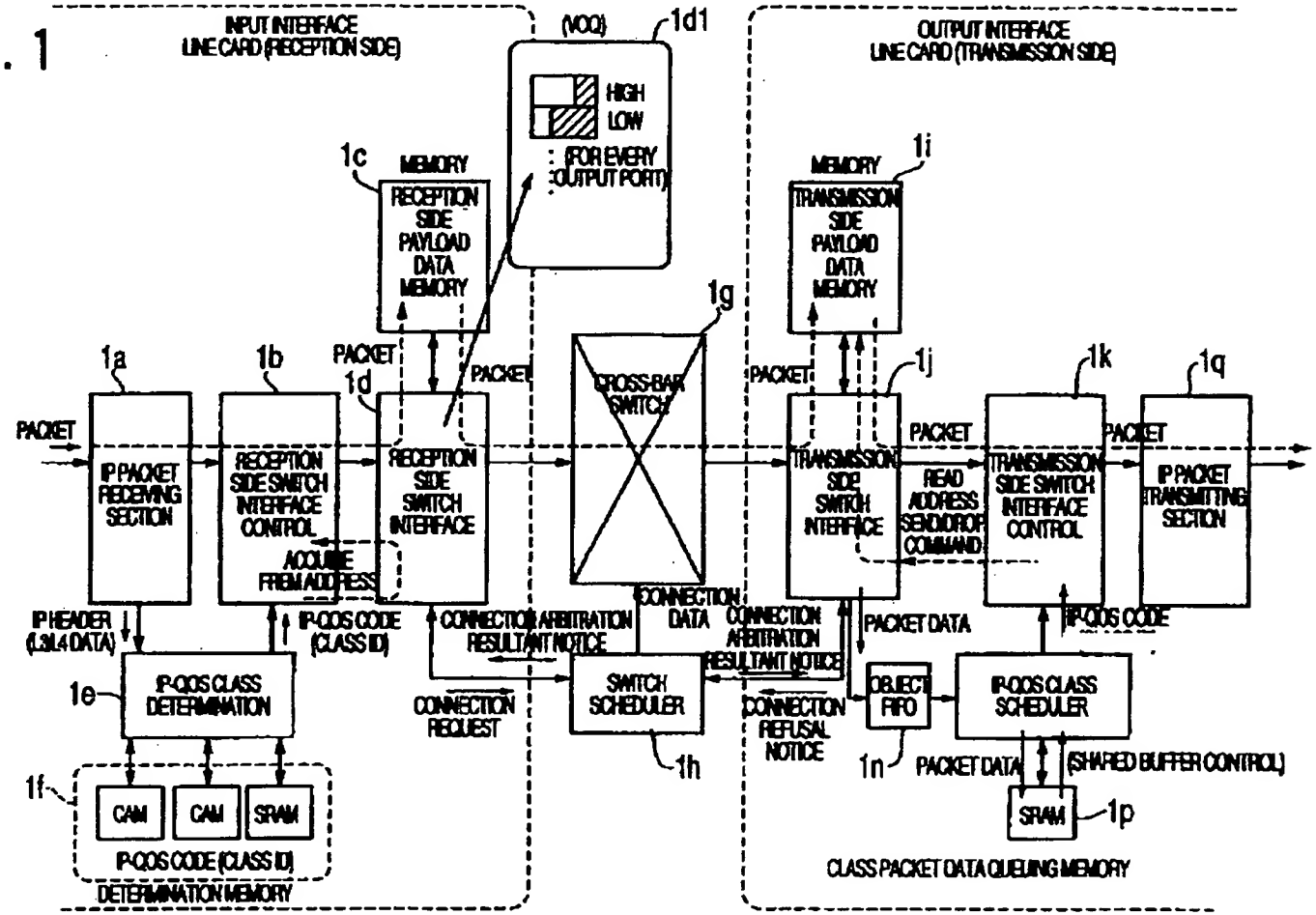

INPUT INTERFACE LINE CARD (RECEPTION SIDE)
(VOQ)
1d1
HIGH
LOW
: (FOR EVERY OUTPUT PORT)
1c
MEMORY
RECEPTION SIDE PAYLOAD DATA MEMORY
1g
CROSS-BAR SWITCH
OUTPUT INTERFACE LINE CARD (TRANSMISSION SIDE)
MEMORY
1i
TRANSMISSION SIDE PAYLOAD DATA MEMORY
1a
1b
1d
PACKET
IP PACKET RECEIVING SECTION
RECEPTION SIDE SWITCH INTERFACE CONTROL
RECEPTION SIDE SWITCH INTERFACE
ACQUIRE FROM ADDRESS
TRANSMISSION SIDE SWITCH INTERFACE
1j
1k
1q
READ ADDRESS SEND/DROP COMMAND
TRANSMISSION SIDE SWITCH INTERFACE CONTROL
IP PACKET TRANSMITTING SECTION
IP HEADER (L3/L4 DATA)
IP-QOS CODE (CLASS ID)
CONNECTION ARBITRATION RESULTANT NOTICE
CONNECTION DATA
PACKET DATA
IP-QOS CODE
1e
IP-QOS CLASS DETERMINATION
CONNECTION REQUEST
SWITCH SCHEDULER
1h
CONNECTION REFUSAL NOTICE
OBJECT FIFO
1n
IP-QOS CLASS SCHEDULER
PACKET DATA
(SHARED BUFFER CONTROL)
1f
CAM
CAM
SRAM
IP-QOS CODE (CLASS ID) DETERMINATION MEMORY
SRAM
1p
CLASS PACKET DATA QUEUING MEMORY

Fig. 2

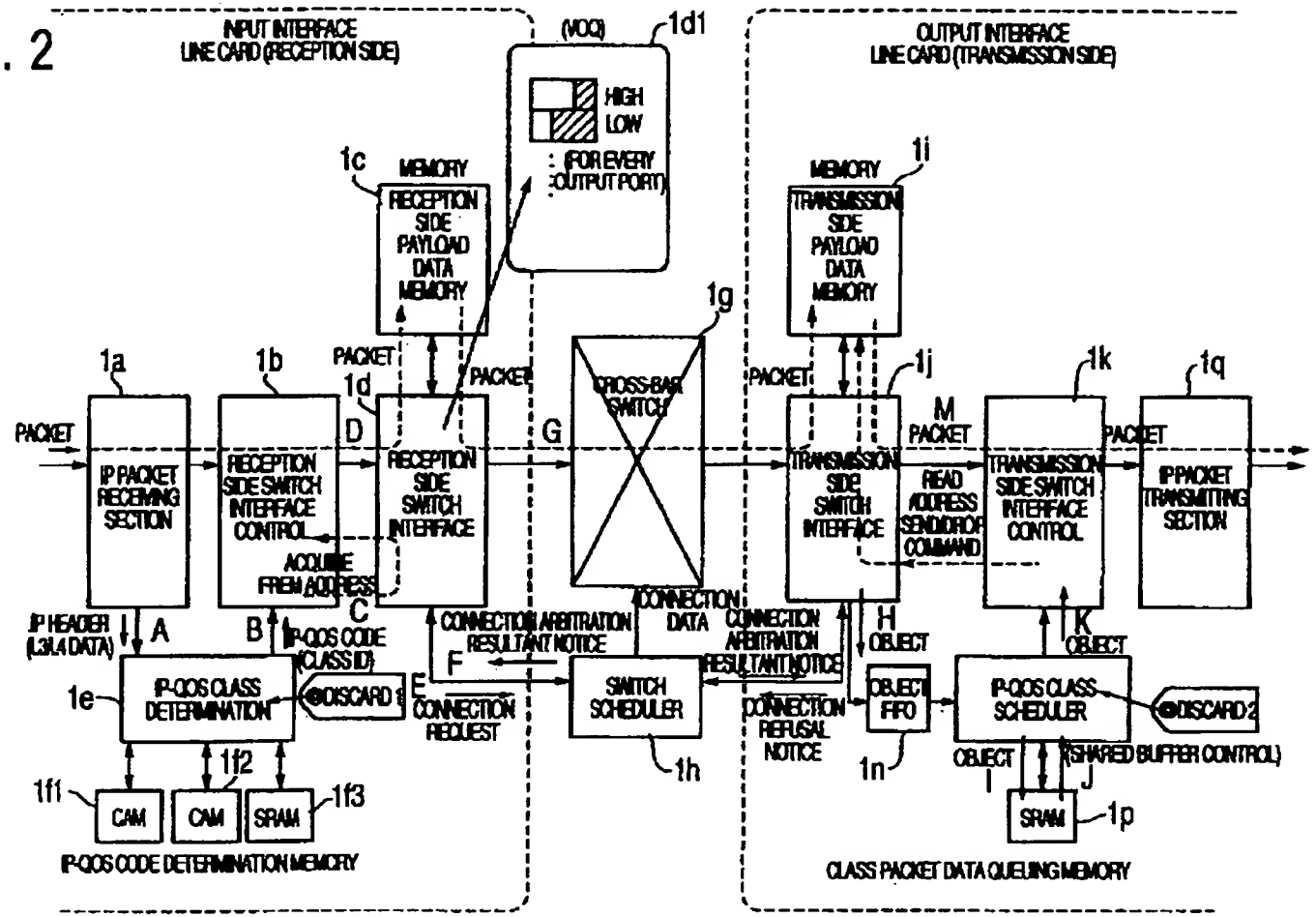

INPUT INTERFACE LINE CARD (RECEPTION SIDE)
(VOQ)
1d1
HIGH
LOW
(FOR EVERY OUTPUT PORT)
OUTPUT INTERFACE LINE CARD (TRANSMISSION SIDE)
1c
MEMORY
RECEPTION SIDE PAYLOAD DATA MEMORY
1i
MEMORY
TRANSMISSION SIDE PAYLOAD DATA MEMORY
1g
CROSS-BAR SWITCH
1a
1b
1d
1j
1k
1q
PACKET
IP PACKET RECEIVING SECTION
RECEPTION SIDE SWITCH INTERFACE CONTROL
RECEPTION SIDE SWITCH INTERFACE
TRANSMISSION SIDE SWITCH INTERFACE
READ ADDRESS SENDING FOR COMMAND
TRANSMISSION SIDE SWITCH INTERFACE CONTROL
IP PACKET TRANSMITTING SECTION
ACQUIRE FREE ADDRESS
D
G
M
A
B
C
H
K
IP HEADER (L3/L4 DATA)
IP-QOS CODE (CLASS ID)
CONNECTION ARBITRATION RESULTANT NOTICE
CONNECTION DATA
OBJECT
1e
IP-QOS CLASS DETERMINATION
DISCARD 1
E
F
CONNECTION REQUEST
SWITCH SCHEDULER
1h
CONNECTION REFUSAL NOTICE
OBJECT FIFO
1n
IP-QOS CLASS SCHEDULER
DISCARD 2
OBJECT
I
J
SHARED BUFFER CONTROL
1f1
1f2
1f3
CAM
CAM
SRAM
1p
IP-QOS CODE DETERMINATION MEMORY
CLASS PACKET DATA QUEUING MEMORY

Fig. 3

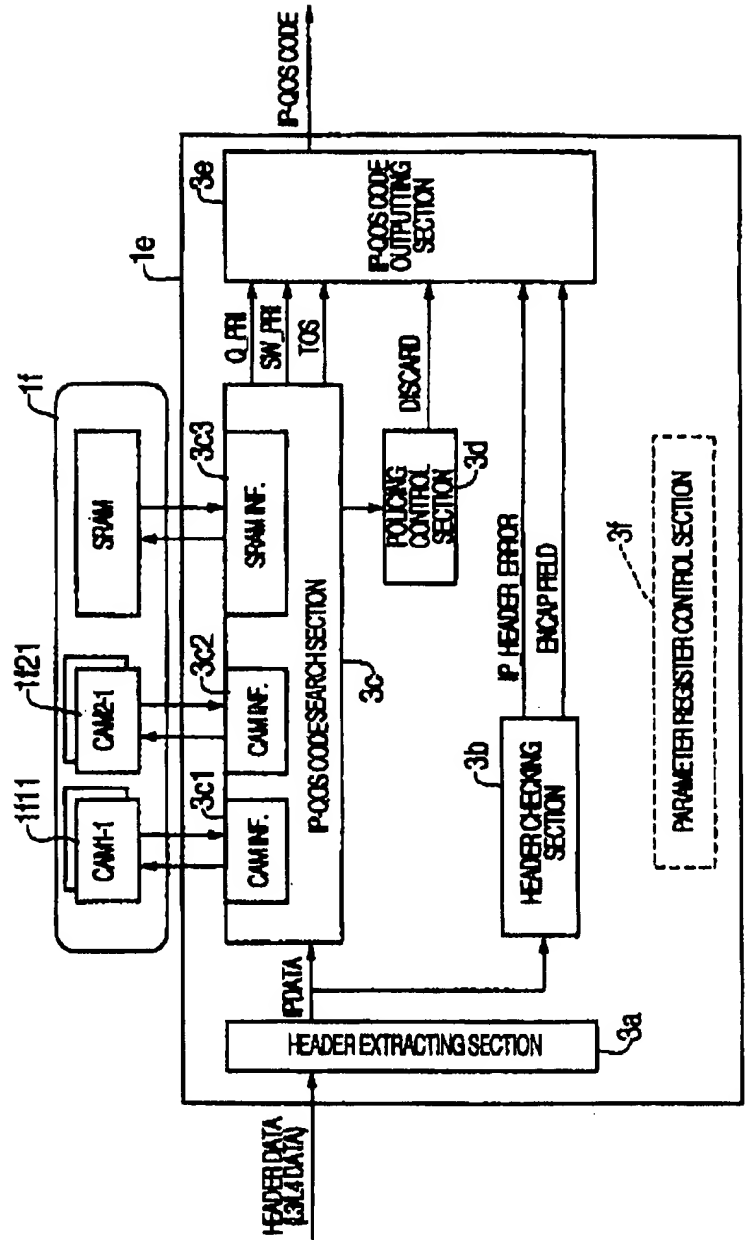

HEADER DATA (SA,DA,DATA)
IP-QOS CODE
1e
3e
IP-QOS CODE OUTPUTTING SECTION
Q_PRI
SW_PRI
TOS
DISCARD
1f
SPAM
3c3
SPAM INF.
POLICING CONTROL SECTION
3d
1f21
CAM2-1
3c2
CAM INF.
IP-QOS CODE SEARCH SECTION
3c
IP HEADER ERROR
ENCAP FIELD
3f
1f11
CAM1-1
3c1
CAM INF.
3b
HEADER CHECKING SECTION
PARAMETER REGISTER CONTROL SECTION
IP DATA
HEADER EXTRACTING SECTION
3a

Fig. 4

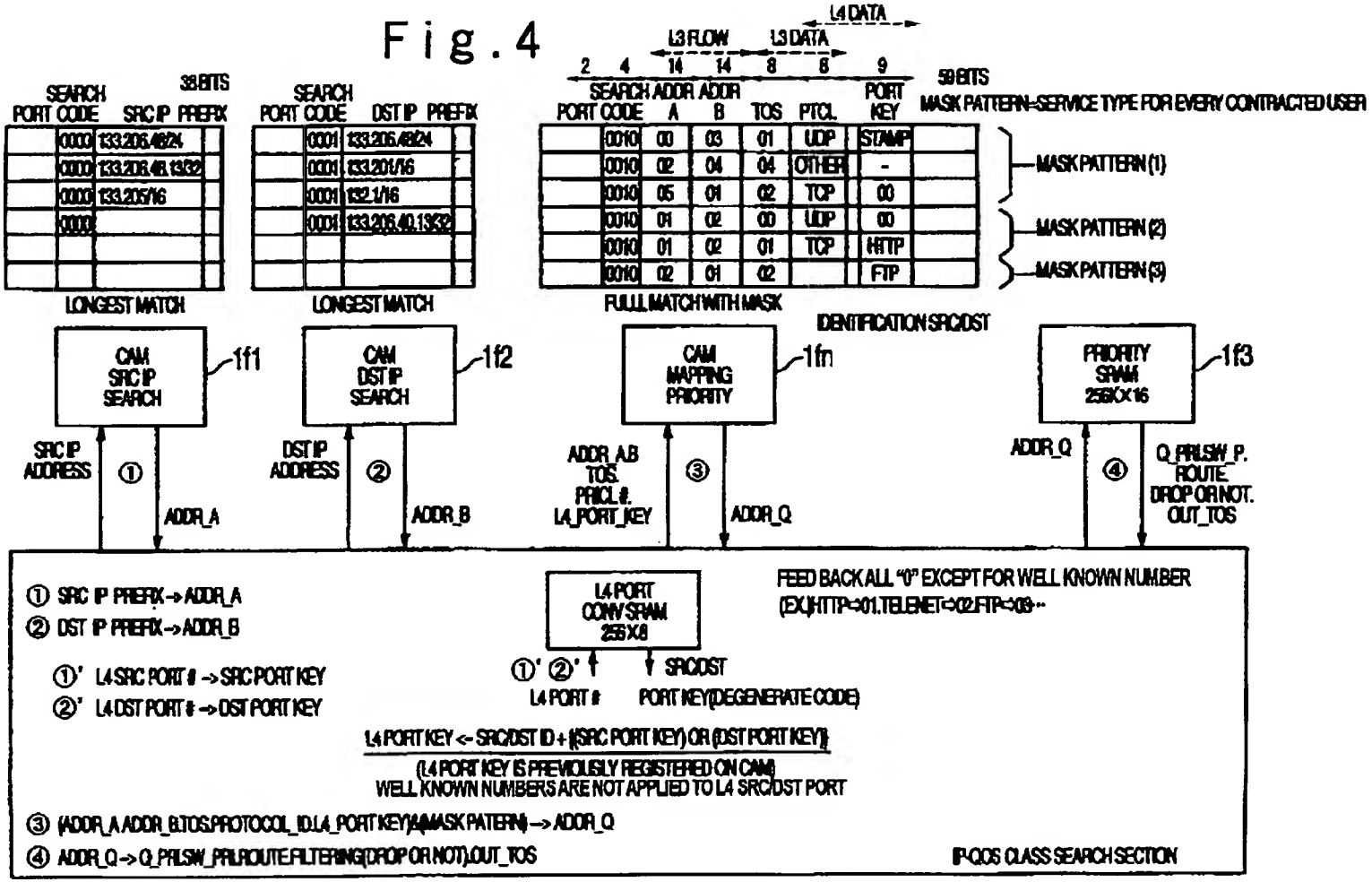

L4 DATA
L3 FLOW
L3 DATA
2 4 14 14 8 8 9
59 BITS
MASK PATTERN=SERVICE TYPE FOR EVERY CONTRACTED USER
38 BITS
SEARCH PORT CODE
SRC IP PREFIX
DST IP PREFIX
ADDR A
ADDR B
TOS
PTCL
PORT KEY
MASK PATTERN (1)
MASK PATTERN (2)
MASK PATTERN (3)
LONGEST MATCH
LONGEST MATCH
FULL MATCH WITH MASK
IDENTIFICATION SRC/DST
CAM SRC IP SEARCH
1f1
CAM DST IP SEARCH
1f2
CAM MAPPING PRIORITY
1fn
PRIORITY SRAM 256X16
1f3
SRC IP ADDRESS
DST IP ADDRESS
ADDR_A B TOS PRICL #. L4_PORT_KEY
ADDR_Q
Q_PRILSW_P. ROUTE DROP OR NOT. OUT_TOS
① SRC IP PREFIX -> ADDR_A
② DST IP PREFIX -> ADDR_B
①' L4 SRC PORT # -> SRC PORT KEY
②' L4 DST PORT # -> DST PORT KEY
L4 PORT CONV SRAM 256X8
①' ②'
L4 PORT #
SRC/DST
PORT KEY(DEGENERATE CODE)
FEED BACK ALL "0" EXCEPT FOR WELL KNOWN NUMBER
(EX.)HTTP=>01.TELENET=>02.FTP=>03···
L4 PORT KEY <- SRC/DST ID + ((SRC PORT KEY) OR (DST PORT KEY))
(L4 PORT KEY IS PREVIOUSLY REGISTERED ON CAM)
WELL KNOWN NUMBERS ARE NOT APPLIED TO L4 SRC/DST PORT
③ (ADDR_A ADDR_B TOS PROTOCOL_ID L4_PORT KEY) && (MASK PATERN) -> ADDR_Q
④ ADDR_Q -> Q_PRILSW_PRI ROUTE FILTERING (DROP OR NOT) OUT_TOS
IP-QOS CLASS SEARCH SECTION

Fig. 5

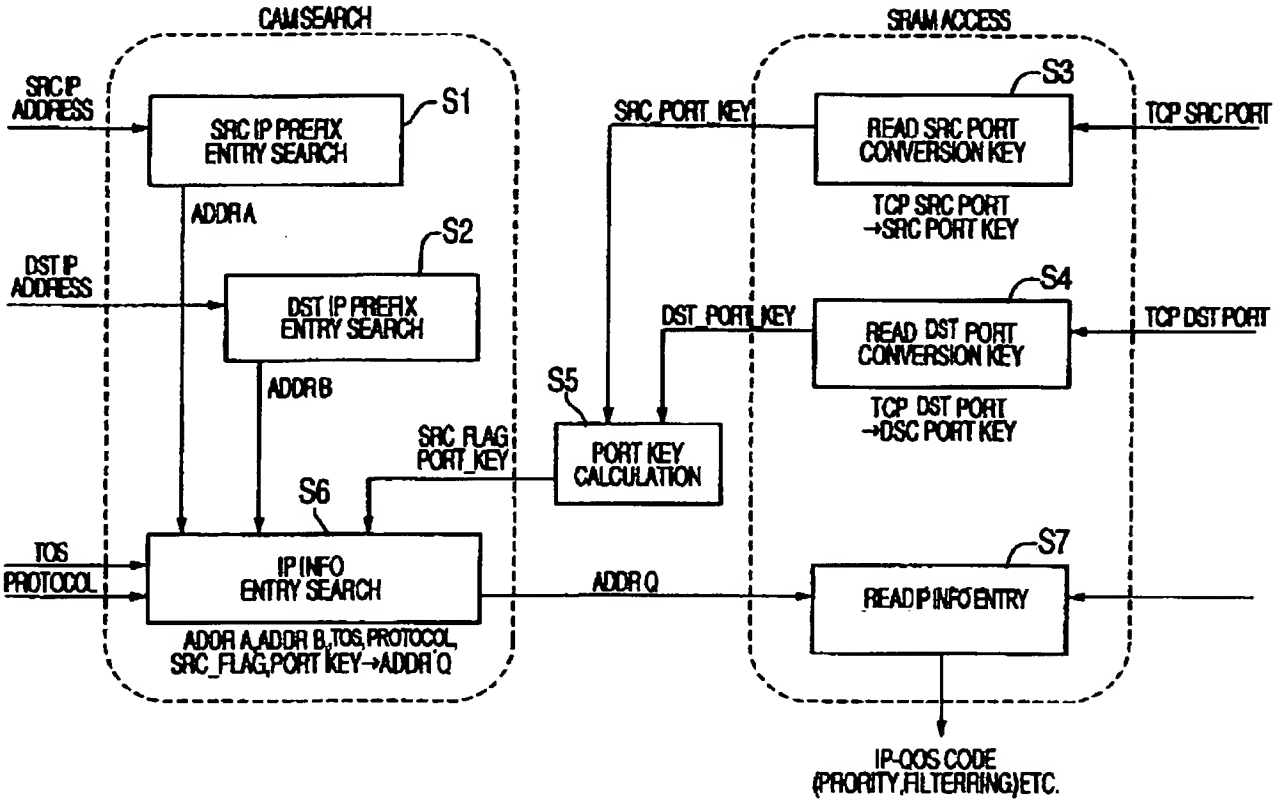

CAM SEARCH
SRAM ACCESS
SRC IP ADDRESS
SRC IP PREFIX ENTRY SEARCH
S1
ADDR A
DST IP ADDRESS
DST IP PREFIX ENTRY SEARCH
S2
ADDR B
S3
SRC_PORT_KEY
READ SRC PORT CONVERSION KEY
TCP SRC PORT
TCP SRC PORT →SRC PORT KEY
S4
DST_PORT_KEY
READ DST PORT CONVERSION KEY
TCP DST PORT
TCP DST PORT →DSC PORT KEY
S5
SRC_FLAG PORT_KEY
PORT KEY CALCULATION
S6
TOS
PROTOCOL
IP INFO ENTRY SEARCH
ADDR A,ADDR B,TOS,PROTOCOL, SRC_FLAG,PORT KEY→ADDR Q
ADDR Q
S7
READ IP INFO ENTRY
IP-QOS CODE (PRIORITY,FILTERRING)ETC.

Fig. 6

【CAM REGION DIVISION】

CAM ADDRESS	CAM DATA (MAX.64 BITS)	MASK PATTERN (64 BITS)	SEARCH METHOD
ADDR_A~	IP SRC PREFIX ENTRY STORAGE REGION		LONGEST MATCH
ADDR_B~	IP DST PREFIX ENTRY STORAGE REGION		LONGEST MATCH
ADDR_Q~	IP INFO SEARCH ENTRY STORAGE REGION		FULL MATCH WITH MASK

【1,IP SRC PREFIX ENTRY STORAGE REGION : SEARCH CODE 0000】

CAM ADDRESS (ADDR_A)	CAM DATA (38 BITS)			
	HW #(2)	SEARCH CODE(4)	IP SRC ADRESS/ PREFIX(32BITS)	NON USED(26 BITS)
A #1	00	0000	IP SRC ADDRESS #1/PREFIX	
A #2	00	0000	IP SRC ADDRESS #2/PREFIX	
A #3	01	0000	IP SRC ADDRESS #1/PREFIX	
⋮	⋮	⋮	⋮	

Fig. 7

【2,IP DST PREFIX ENTRY STORAGE REGION : SEARCH CODE 0001】

CAM ADDRESS (ADDR_B)	CAM DATA (38 BITS)			
	HW #(2)	SEARCH CODE(4)	IP DST ADDRESS/ PREFIX(32BITS)	NON USED(26 BITS)
B #1	00	0001	IP DST ADDRESS #1/PREFIX	
B #2	00	0001	IP DST ADDRESS #2/PREFIX	
B #3	01	0001	IP DST ADDRESS #1/PREFIX	
⋮	⋮	⋮	⋮	

【3,IP INFO ADDRESS ENTRY STORAGE REGION : SEARCH CODE 0010】

CAM ADDRESS (ADDR_Q)	CAM DATA (55 BITS)								
	HW #(2)	SEARCH CODE(4)	ADDR A(14)	ADDR B(14)	TOS (8)	PROTO COL# (8)	SRC/ DST (1)	PORT KEY (8)	NON USED (5 BITS)
Q #1	00	0010	A1	B1	01	TCP	S	HTTP	
Q #2	00	0010	A1	B2	04	UDP	D	SNMP	
Q #3	01	0010	A3	B1	02	TCP	S	FTP	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	

Fig.8

[IP INFO ENTRY]

ADDRESS(16 BITS) : UPPER 2 BITS=00 LOWER 14 BITS=HIT ADDR_Q	DATA(24 BIT)							
	Q_PRI(4)	D	P	ROUTE(1+4)		OUTPUT TOS(2+6)		RESERVE (3)
ADDR Q0	0000	0	0	0	0000	11	011011 00	
ADDR Q1	1101	0	1	0	0000	11	011010 00	
ADDR Q2	1101	0	0	0	0000	00	000000 00	
⋮	⋮	⋮	⋮		⋮		⋮	
ADDR QI	1110	0	1	1	0101	00	000000 00	
⋮	⋮	⋮	⋮		⋮		⋮	

Fig. 9

(IPV4 & TCP/UDP/OTHER HEADER FORMAT)

WORD	63			47	31			15
-	EMPTY DATA				PPP HEADER			
0	VER	IHL	TOS	DETAGRAM LENGTH	IDENTIFICATION		M	FRAGMENT OFFSET
1	TTL		PROTOCOL	HEADER CHECKSUM	SRC IP ADDRESS			
2	DST IP ADDRESS				L4 SRC PORT			L4 DST PORT

Fig. 10

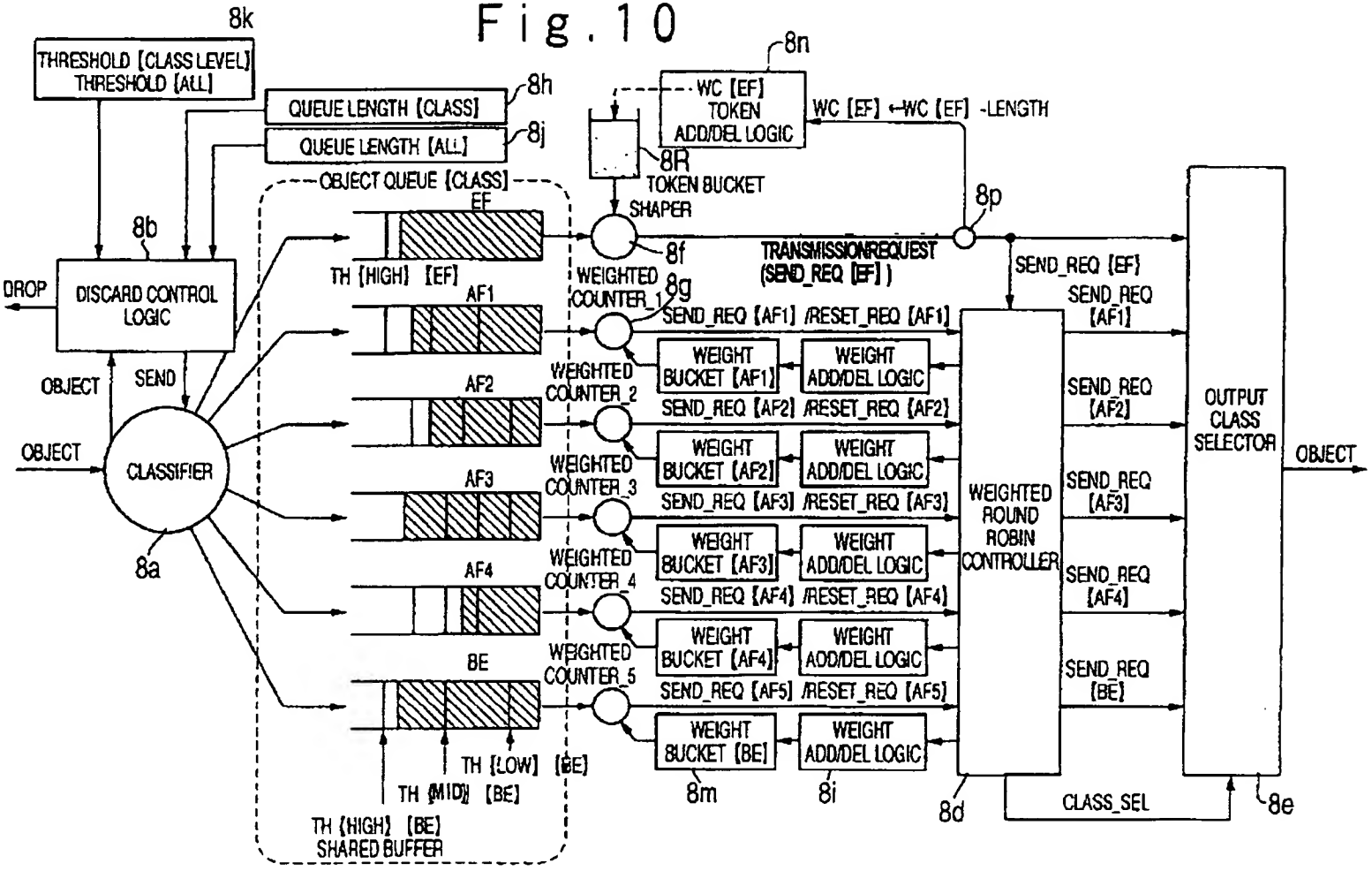

8k
THRESHOLD [CLASS LEVEL]
THRESHOLD [ALL]
QUEUE LENGTH [CLASS]
8h
QUEUE LENGTH [ALL]
8j
8n
WC [EF] TOKEN ADD/DEL LOGIC
WC [EF] ←WC [EF] -LENGTH
8R
TOKEN BUCKET
SHAPER
8p
OBJECT QUEUE [CLASS]
EF
8b
8f
TRANSMISSIONREQUEST (SEND_REQ [EF])
SEND_REQ [EF]
DROP
DISCARD CONTROL LOGIC
TH [HIGH] [EF]
WEIGHTED COUNTER_1
8g
AF1
SEND_REQ [AF1] /RESET_REQ [AF1]
WEIGHT BUCKET [AF1]
WEIGHT ADD/DEL LOGIC
SEND_REQ [AF1]
OBJECT
SEND
AF2
WEIGHTED COUNTER_2
SEND_REQ [AF2] /RESET_REQ [AF2]
WEIGHT BUCKET [AF2]
WEIGHT ADD/DEL LOGIC
SEND_REQ [AF2]
OUTPUT CLASS SELECTOR
OBJECT
OBJECT
CLASSIFIER
AF3
WEIGHTED COUNTER_3
SEND_REQ [AF3] /RESET_REQ [AF3]
WEIGHT BUCKET [AF3]
WEIGHT ADD/DEL LOGIC
WEIGHTED ROUND ROBIN CONTROLLER
SEND_REQ [AF3]
8a
AF4
WEIGHTED COUNTER_4
SEND_REQ [AF4] /RESET_REQ [AF4]
WEIGHT BUCKET [AF4]
WEIGHT ADD/DEL LOGIC
SEND_REQ [AF4]
BE
WEIGHTED COUNTER_5
SEND_REQ [AF5] /RESET_REQ [AF5]
WEIGHT BUCKET [BE]
WEIGHT ADD/DEL LOGIC
SEND_REQ [BE]
TH [LOW] [BE]
TH [MID] [BE]
8m
8i
8d
CLASS_SEL
8e
TH [HIGH] [BE]
SHARED BUFFER

Fig. 11

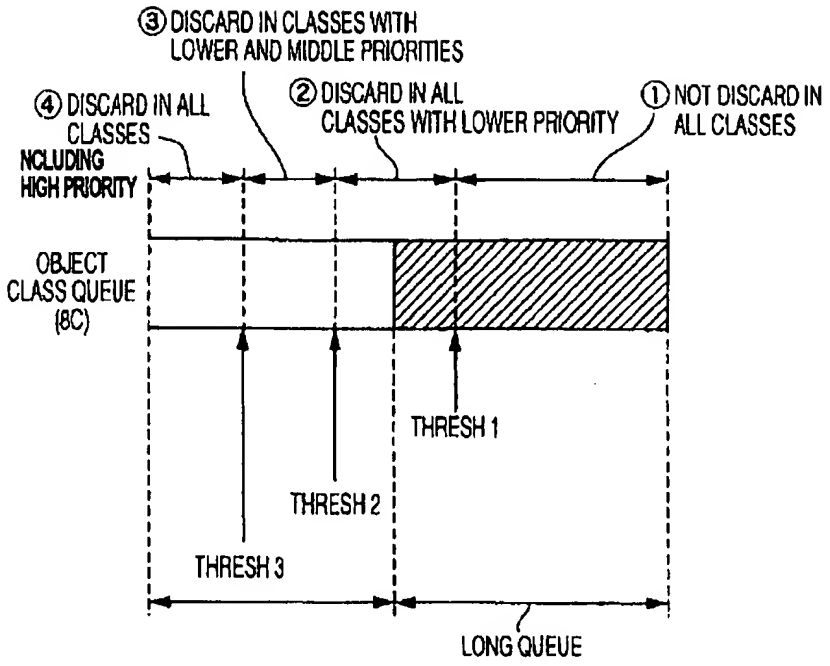

③ DISCARD IN CLASSES WITH LOWER AND MIDDLE PRIORITIES
④ DISCARD IN ALL CLASSES
② DISCARD IN ALL CLASSES WITH LOWER PRIORITY
① NOT DISCARD IN ALL CLASSES
NCLUDING HIGH PRIORITY
OBJECT CLASS QUEUE (8C)
THRESH 1
THRESH 2
THRESH 3
LONG QUEUE

Fig. 12

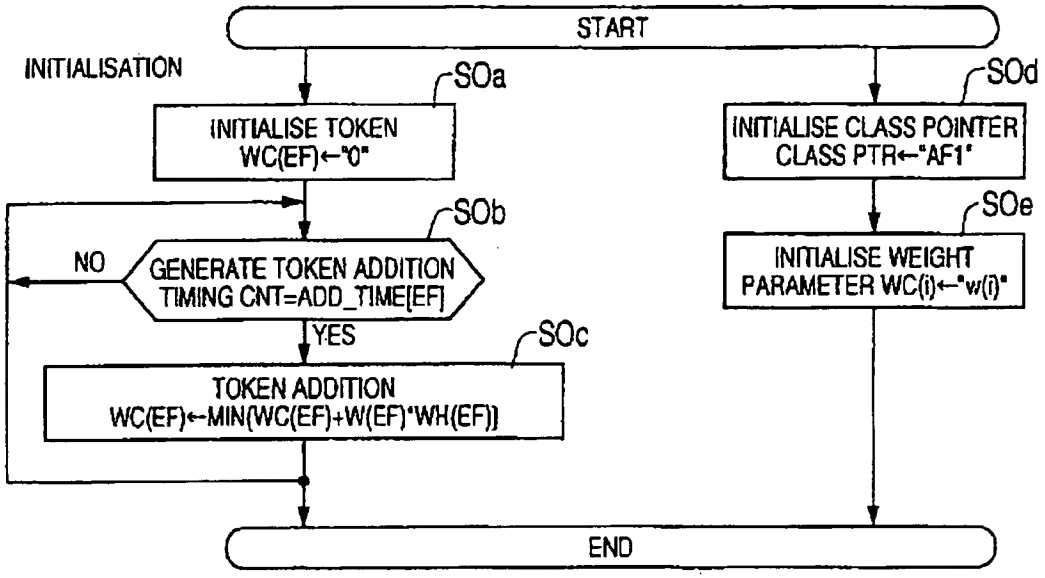

START
INITIALISATION
S0a
INITIALISE TOKEN
WC(EF)←"0"
S0b
NO
GENERATE TOKEN ADDITION
TIMING CNT=ADD_TIME[EF]
YES
S0c
TOKEN ADDITION
WC(EF)←MIN(WC(EF)+W(EF)*WH(EF))
S0d
INITIALISE CLASS POINTER
CLASS PTR←"AF1"
S0e
INITIALISE WEIGHT
PARAMETER WC(i)←"w(i)"
END

Fig. 13

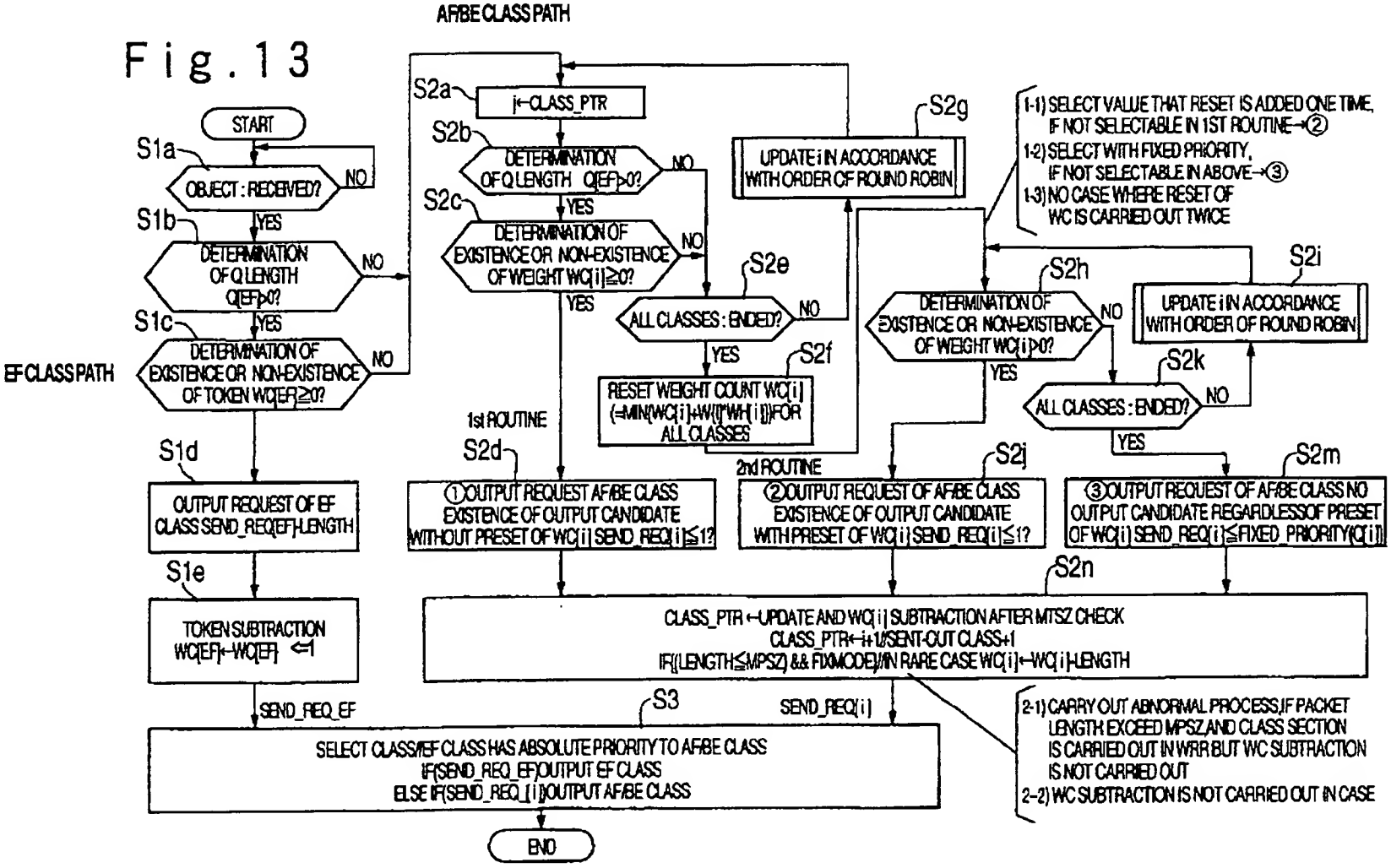

AF/BE CLASS PATH
EF CLASS PATH
START
S1a
OBJECT : RECEIVED?
NO
YES
S1b
DETERMINATION OF Q LENGTH Q[EF]>0?
NO
YES
S1c
DETERMINATION OF EXISTENCE OR NON-EXISTENCE OF TOKEN WC[EF]≧0?
NO
S1d
OUTPUT REQUEST OF EF CLASS SEND_REQ[EF]=LENGTH
S1e
TOKEN SUBTRACTION WC[EF]←WC[EF] ⇐1
SEND_REQ_EF
S2a
j←CLASS_PTR
S2b
DETERMINATION OF Q LENGTH Q[EF]>0?
NO
YES
S2c
DETERMINATION OF EXISTENCE OR NON-EXISTENCE OF WEIGHT WC[i]≧0?
NO
YES
S2e
ALL CLASSES : ENDED?
NO
YES
S2f
RESET WEIGHT COUNT WC[i] (=MIN{WC[i]+W([i]WH[i])}FOR ALL CLASSES
1st ROUTINE
S2d
①OUTPUT REQUEST AF/BE CLASS EXISTENCE OF OUTPUT CANDIDATE WITHOUT PRESET OF WC[i] SEND_REQ[i]≦1?
2nd ROUTINE
S2g
UPDATE i IN ACCORDANCE WITH ORDER OF ROUND ROBIN
S2h
DETERMINATION OF EXISTENCE OR NON-EXISTENCE OF WEIGHT WC[i]>0?
NO
YES
S2i
UPDATE i IN ACCORDANCE WITH ORDER OF ROUND ROBIN
S2k
ALL CLASSES : ENDED?
NO
YES
S2j
②OUTPUT REQUEST OF AF/BE CLASS EXISTENCE OF OUTPUT CANDIDATE WITH PRESET OF WC[i] SEND_REQ[i]≦1?
S2m
③OUTPUT REQUEST OF AF/BE CLASS NO OUTPUT CANDIDATE REGARDLESS OF PRESET OF WC[i] SEND_REQ[i]≦FIXED_PRIORITY(YC[i])
1-1) SELECT VALUE THAT RESET IS ADDED ONE TIME, IF NOT SELECTABLE IN 1ST ROUTINE→②
1-2) SELECT WITH FIXED PRIORITY, IF NOT SELECTABLE IN ABOVE→③
1-3) NO CASE WHERE RESET OF WC IS CARRIED OUT TWICE
S2n
CLASS_PTR←UPDATE AND WC[i] SUBTRACTION AFTER MTSZ CHECK
CLASS_PTR←i+1//SENT-OUT CLASS+1
IF{(LENGTH≦MPSZ) && FIXMODE}//IN RARE CASE WC[i]←WC[i]-LENGTH
SEND_REQ[i]
2-1) CARRY OUT ABNORMAL PROCESS IF PACKET LENGTH EXCEED MPSZ AND CLASS SECTION IS CARRIED OUT IN WRR BUT WC SUBTRACTION IS NOT CARRIED OUT
2-2) WC SUBTRACTION IS NOT CARRIED OUT IN CASE
S3
SELECT CLASS//EF CLASS HAS ABSOLUTE PRIORITY TO AF/BE CLASS
IF{SEND_REQ_EF}OUTPUT EF CLASS
ELSE IF{SEND_REQ_[i]}OUTPUT AF/BE CLASS
END

Fig. 14

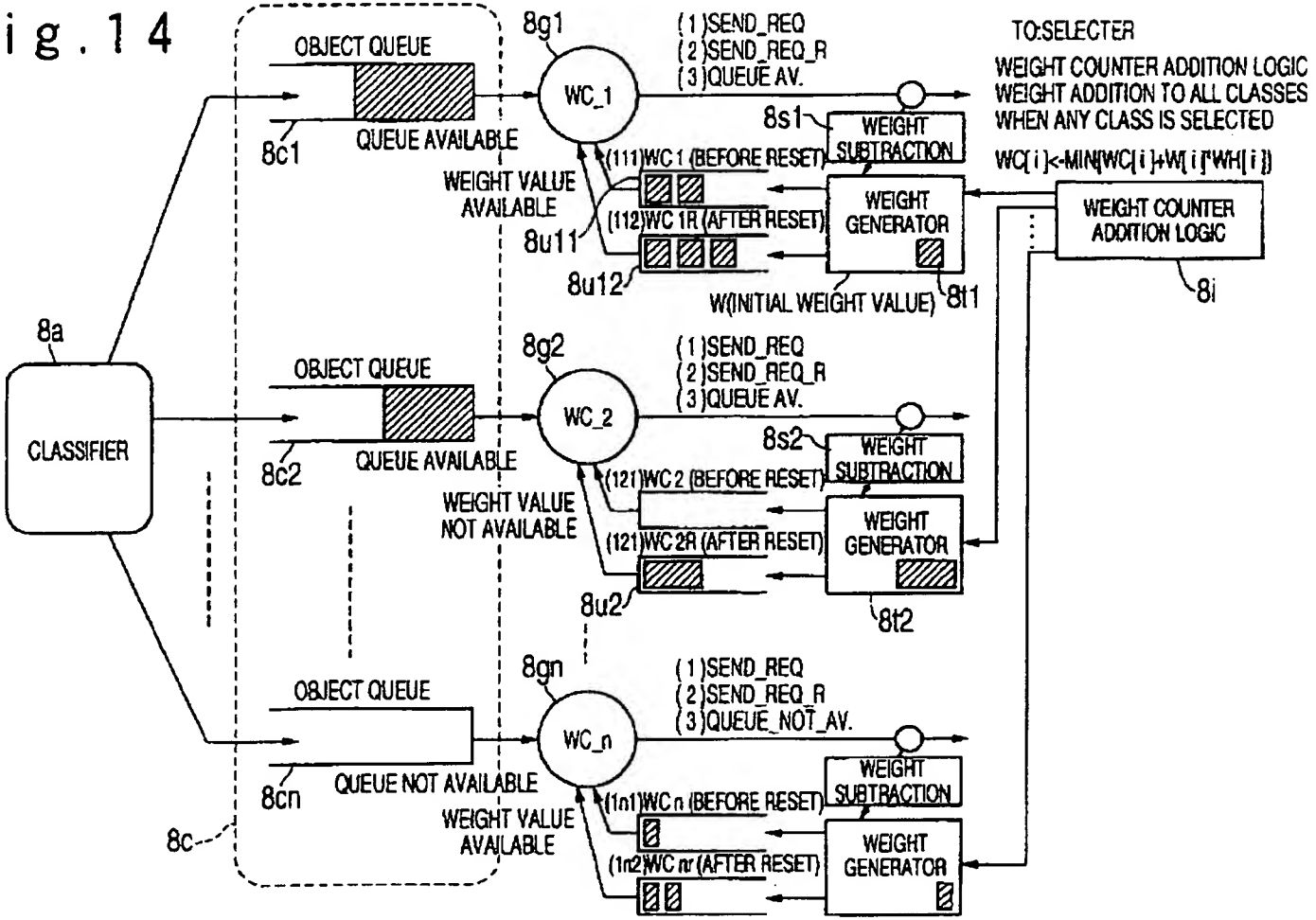

8a
CLASSIFIER
OBJECT QUEUE
8c1
QUEUE AVAILABLE
WEIGHT VALUE AVAILABLE
8g1
WC_1
(1)SEND_REQ
(2)SEND_REQ_R
(3)QUEUE AV.
8s1
WEIGHT SUBTRACTION
(111)WC 1 (BEFORE RESET)
(112)WC 1R (AFTER RESET)
8u11
8u12
WEIGHT GENERATOR
W(INITIAL WEIGHT VALUE)
8t1
TO:SELECTER
WEIGHT COUNTER ADDITION LOGIC
WEIGHT ADDITION TO ALL CLASSES WHEN ANY CLASS IS SELECTED
WC[i]<-MIN[WC[i]+W[i]*WH[i]]
WEIGHT COUNTER ADDITION LOGIC
8i
OBJECT QUEUE
8c2
QUEUE AVAILABLE
WEIGHT VALUE NOT AVAILABLE
8g2
WC_2
(1)SEND_REQ
(2)SEND_REQ_R
(3)QUEUE AV.
8s2
WEIGHT SUBTRACTION
(121)WC 2 (BEFORE RESET)
(121)WC 2R (AFTER RESET)
8u2
WEIGHT GENERATOR
8t2
OBJECT QUEUE
8cn
QUEUE NOT AVAILABLE
WEIGHT VALUE AVAILABLE
8c
8gn
WC_n
(1)SEND_REQ
(2)SEND_REQ_R
(3)QUEUE_NOT_AV.
WEIGHT SUBTRACTION
(1n1)WC n (BEFORE RESET)
(1n2)WC nr (AFTER RESET)
WEIGHT GENERATOR

Fig. 15

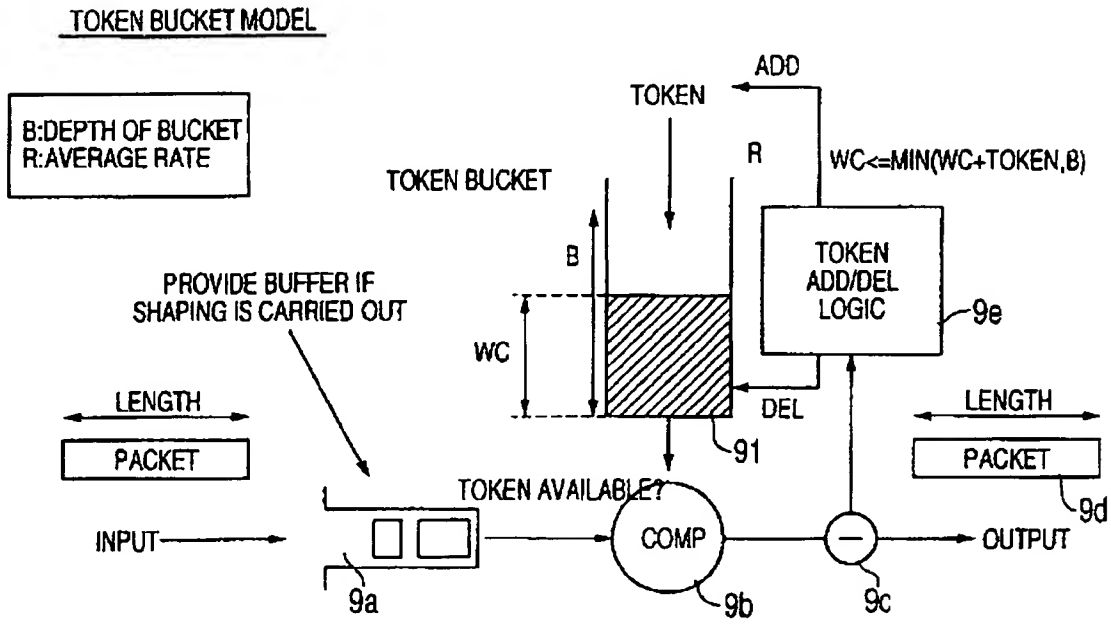
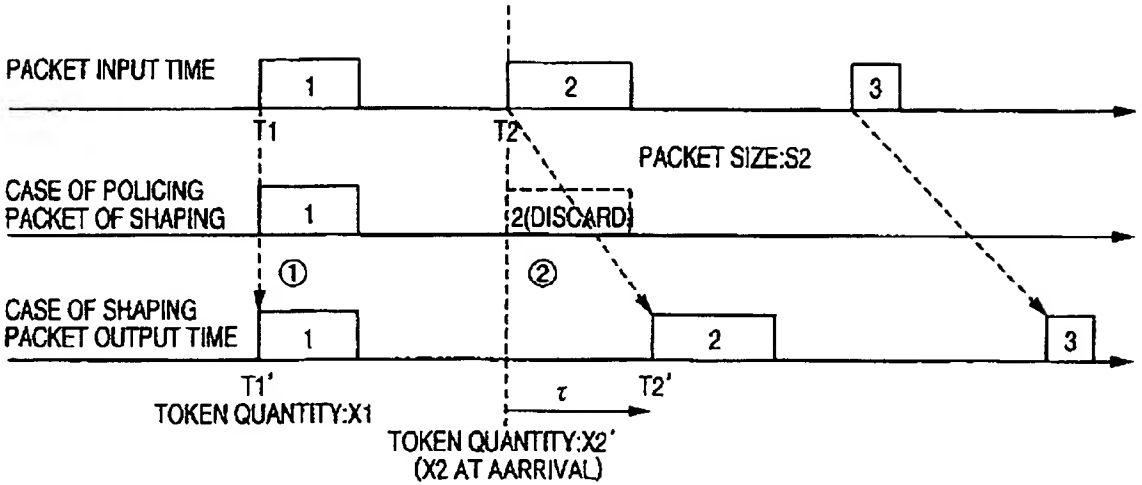

TOKEN BUCKET MODEL
B:DEPTH OF BUCKET
R:AVERAGE RATE
TOKEN
ADD
R
WC<=MIN(WC+TOKEN,B)
TOKEN BUCKET
B
TOKEN ADD/DEL LOGIC
9e
PROVIDE BUFFER IF SHAPING IS CARRIED OUT
WC
LENGTH
PACKET
DEL
91
LENGTH
PACKET
9d
TOKEN AVAILABLE?
INPUT
COMP
OUTPUT
9a
9b
9c

Fig. 16



1)TOKEN QUANTITIY AT TIME T2:X2=x1+(T2-T1)*R
LACK OF TOKEN,IF S2>X2
2)POLICING
IMMEDIATELY DISCARD
3)SHAPING
NOT LACK OF TOKEN,IF PACKET IS TRANSMITTED AT TIME
(τ+T2)(S2=X1+{(τ+T2)-T1}*R).PACKET IS TRANSMITTED WITH DELAY τ